Paper I

Principles of Multivariate Image Analysis (MIA) in remote sensing, technology and industry

Name:	Kim H. Esbensen[*#]	Thorbjørn T. Lied[*]	Kim Lowell[#]	Geoffrey Edwards[#]
e-mail:	Kim.Esbensen@hit.no	thorbjorn@lied.no	Kim.Lowell@scg.ulaval.ca	Geoffrey.edwards@geoide.ulaval.ca
Phone:	+47 35 57 51 50	+ 47 35 57 51 53	+1 (418) 656-5491	+1 (418) 656-2131
Fax:	+47 35 57 52 50	+ 47 35 57 52 50	+1 (418) 656-3607	+1 (418) 656-3607

(*) Applied Chemometrics Research Group (ACRG),

Telemark University College, Institute of Technology (HIT-TF), Porsgrunn, Norway

(#) Centre de Recherché en géomatique

Université Laval, Québec, Qc G1K 7P4, Canada

Principles of Multivariate Image Analysis (MIA) in remote sensing, technology and industry

ABSTRACT

Multivariate Image Analysis (**MIA**) is a set of *interdependent image analysis procedures*, encompassing both an explorative and pattern recognition mode, as well as classification-related facilities. A synoptic overview of all elements in MIA is presented by a series of different applications on a comparison data set, from Montmorency Forest, Quebec (a high resolution, airborne spectrophotometric image).

We present extensive justifications for the specific MIA approach - *contra* a more traditionalist image processing mode. MIA is specifically **not** a one-shot image analytical method, but a set of optional, interdependent explorative analysis and classification tools to be used by the *informed* analyst, according to the *specific objectives* of the image analysis *problem context*. This overview will furnish the image analyst with an overview of the *principles of MIA* needed to compose one's own flexible, problem-specific strategy of multivariate image analysis: pixel class delineation in feature space (*score* space) by a topographic analogy and an end-member/mixing class concept.

The exposition forms a user-oriented *complement* to the introductory book: "Multivariate Image Analysis" (Geladi and Grahn 1996).

INTRODUCTION

Esbensen and Geladi (1989) and Lowell and Esbensen (1993) argued at length against what was termed a "traditionalist" image analysis paradigm, more specifically the universally adopted tradition of starting any image analysis in scene space. It was argued that this necessarily must lead to *sub-optimal class representations* amongst other deficiencies. This traditional approach is centred upon the concept of delineating scene-space areas (objects, or part-object) that are as homogeneous and/or spatially coherent as possible, in order to find "representative" training classes. It was shown that this is but an optimistic *hope* at best. In addition there often would appear to be a slight confusion in this image analysis tradition by not always specifying clearly in advance whether one is engaged in unsupervised (exploratory) or supervised (discrimination/classification) undertakings; see in-depth discussion by Lowell and Esbensen (1993). We shall here add only a few, but salient additional iconoclastic comments along the way of presenting the MIA concepts and illustrations below.

MIA approach: synopsis

MIA takes its point of departure in feature space in general, in the so-called score space in particular (Esbensen and Geladi 1989, Geladi and Grahn 1996). MIA can to a first delineation be understood as a *truncated* principal component modelling of the multi-channel image, producing sets of complementary score- and loading-plots. It was argued that MIA's main thrust is that the score-plot comprise the *only valid starting point for any image analysis*, in that this is the only completely comprehensive

delineation of the channel covariance structure(s) of the entire image pixel aggregation. The score-plot visualises the entire image inter-pixel dispositions (pixel classes, groupings, trends, outliers i.a.), while the complementary loading plot gives a graphic illustration of the underlying channel covariance/correlations responsible for the score dispositions.

DATA PRESENTATION

Master data set: rationale

The comparison data set to be used for all examples in this overview is an 8-channel digital image, recorded with the Canadian MEIS II airborne spectrometer, with a spatial pixel resolution of 70 cm, as described in McColl et al. (1984) and Esbensen at al. (1993). While in satellite-based remotely sensed satellite imagery usually all fine textural detail is lost at most of the presently available resolutions, digital analysis of the present type of very high resolution airborne imagery of forested scenes have also been fraught with difficulties, but here because the data is so highly variable that traditional pixel classifiers have been more or less ineffective. With high spatial resolution, individual image pixels (very) often tend to cover only minute fractions of the image objects and a pronounced "smearing" of discriminable feature space classes is often observed. This is well-known the *mixed pixel problem*.

This particular data set was chosen because it represents an easily manageable dimensionality, 8 channels. While clearly at the low(er) end of what is representative for modern remote sensing, or for technological/industrial imaging spectrometers, say 10 through 256, 512 channels or even more (be they laboratory instruments or otherwise), this dimensionality still allows us to present all the principles and the potential of the MIA approach. Perhaps more important, the data set was chosen because of it's particularly high resolution "overrepresentation" of mixed pixels, but it will serve equally well also in relation to less resolved (as well for as even more densely sampled image types than the present, such as in current medical, tomographic, chemical imaging). It is emphasized however that this data set is primarily **a vehicle for presenting the general MIA approach**, and that the various examples of MIA analyses below should not be interpreted as *specific* for remote sensing type imagery - On the contrary, all our examples and illustrations have been selected because of their relevance to the general MIA application potential.

Montmorency Forest, *Quebec, Canada – forestry background*

This forest scene was acquired in September 1986 over the Montmorency experimental forest belonging to Laval University, Quebec by the Canadian MEIS II airborne platform; MEIS II is described in detail in Kramer (1996) (B.115). Table 1 list the pertinent spectroscopic channel characteristics in the visible and near infrared.

Table 1. MEIS II channel characteristics

Band *Wavelength (nm)* *Bandwidth (nm)*

1	776	37.0
2	675	39.5
3	747	16.7
4	481	30.9
5	734	16.9
6	710	15.6
7	698	13.1
8	549	31.9

The Montmorency Forest is located 80 km north of Quebec City near the southern edge of the Northern boreal forest that dominates much of Canada. This partly heavily terrained forest contains mainly balsam fir, with minor white birch, white spruce and a small number of red spruce, trembling aspen and some other species.

The site selected for this study contains a cutover which was cleared in 1975 and 1978, and which has both natural and planted conifer regeneration resulting from the forest experimental and observation campaigns in the ensuing period through 1986. Fig. 1 shows the master Montmorency Forest scene, as depicted by MEIS-II channels 1:2:7 (R:G:B).

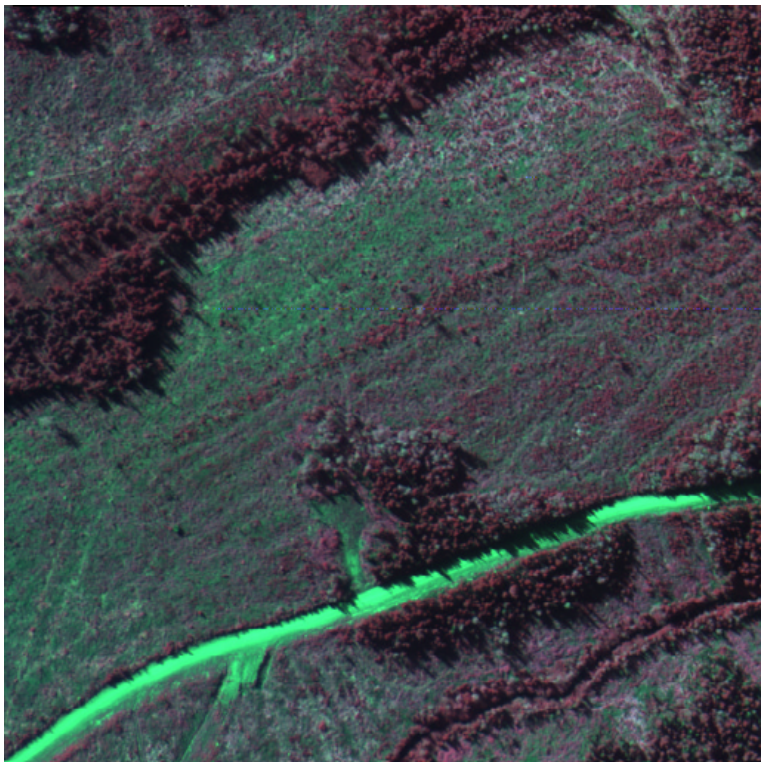


Figure 1. Composite scene display (R/G/B: 1/2/7)

The cut clearing is characterised by balsam fir stands around the perimeter, one white birch stand in the central area, a small stream (bottom right), east-west trending dirt roads and so on. The spatial and temporal regeneration history of this multiple cutover has been the subject of several extensive Laval University Forestry Department studies and is

accordingly very well understood. This scene is called the "Clear-cut Study" in the illustrations that follow below. Further in-depth scene description and full background forestry references i.a. was given by Esbensen et al. (1993). The specific scene history, as will transpire below, turned out to be a particularly illustrative context both for illustrating the comprehensive MIA approach as well as allowing a powerful insight into the possibilities for spatio-temporal analysis of imagery characterised by emergent structures, i.e. temporally and/or spatially evolving/growing/changing structures.

TOOLS IN MIA

MIA score space starting point

Figs 2 & 3 show the most relevant PC-component cross-plots pertaining to the Montmorency Forest scene.

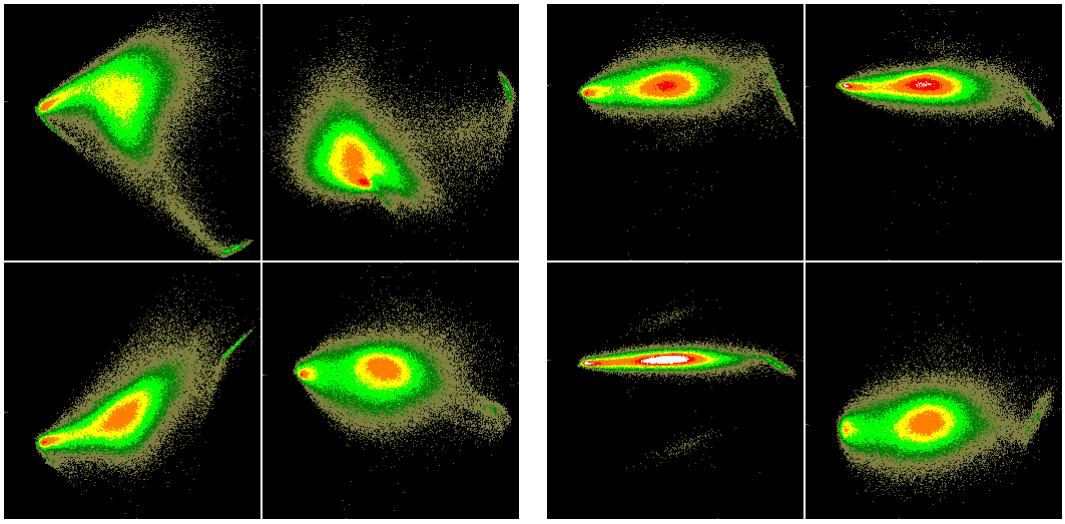


Figure 2. Standard MIA split-screen score plots (PC12,PC13,PC14,PC23)

Note how the main MIA score cross-plots specifically uses the *same PC-component* as the X-axis; this usually is PC-1, although the image analyst may opt for any alternative (e.g. PC-2 or PC-3), because of the well-known decreasing variance associated with an increasing number of principal components calculated. In typical remote sensing imagery and in many other comparative types of imagery, PC-1 often represents an *overall* albedo/reflection/.... *intensity, or contrast measure,* that either may, or may not, be well suited for this common X-axis role depending on the specific image analysis context (hence the alternatives).

Every scene has it's own *distinct* score-space layout - although many similarities and analogies eventually will be noted in building up one's own multivariate image analysis experience. The integrity and individuality of each new multivariate image that is to be analysed cannot be overemphasized. The score space layout cannot be anticipated in advance, hence there is *never* any given *a priori* method ("algorithm") for exploratory

image analysis, - and neither for classification or for segmentation. This is the first main difference between traditional image analysis and MIA.

Every scene has it's own *distinct* score-space layout - although many similarities and analogies eventually will be noted in building up one's own multivariate image analysis experience. The integrity and individuality of each new multivariate image that is to be analysed cannot be overemphasized. The score space layout cannot be anticipated in advance, hence there is *never* any given *a priori* method ("algorithm") for exploratory image analysis, - and neither for classification or for segmentation. This is the first main difference between traditional image analysis and MIA.

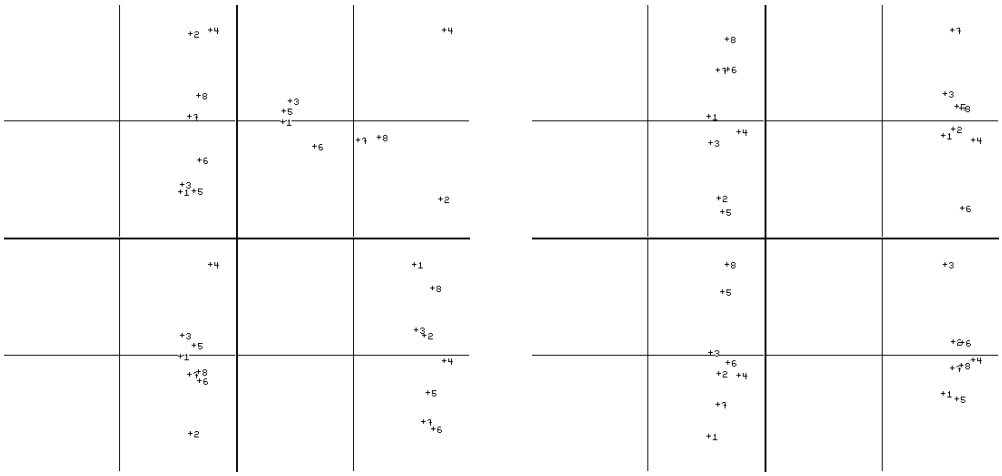


Figure 3. Standard MIA split-screen loading plots (cmp. Fig. 2 for layout)

By way of contrast this paper endeavours to develop a meta-principle for a general approach to multivariate image analysis – and it will be shown that a phenomenological analogy with a topographic map will be very useful.

Fig. 2 shows the entire score space layout (first four principal components) of the Montmorency Forest scene in question (all 512 x 512 pixels are included in the PC-analysis). When a new image is to be analysed, the series of MIA score cross-plot is the only systematic, comprehensive approach to the objective data structures present in the feature domain. This score-plot array will necessarily display all there is to be learned from inspection of the spectral data structure of the image, provided, of course, that all pixels have been included in the analysis. The MIA approach is designed upon this central concept of having access to all image pixels, which is (very nearly) always an easy task with today's PC-power. Without loss of generality we may assume that all pixels are included in the analysis in the expositions below (but even when this cannot be achieved, in some specific hardware configuration case, MIA's design philosophy allows for easy remedies, Esbensen & Geladi, 1989).

The PC-12 score plot always carries the largest fraction of variance modelled (see standard principal components analysis theory) and is consequently always assessed first. MIA analysts should always make due note of the relative proportions of the total variance modelled by each component image. In the present case PC1 and PC2 "explain"

(as this modelling parlance goes) 35% and 27% respectively, totalling 62% of the trace of overall variance in the spectral covariance X'X matrix. The image analyst should always take notice of the individual as well as the accumulated fractions pertaining to all score cross-plots inspected so far, lest interpretations accidentally be based on a too meagre residual variance. For the Montmorency Forest scene the decreasing variance fraction for all eight components breaks down as follows: 35, 27, 10, 9, …. %, which is by and large typical of a large number of multivariate imagery types.

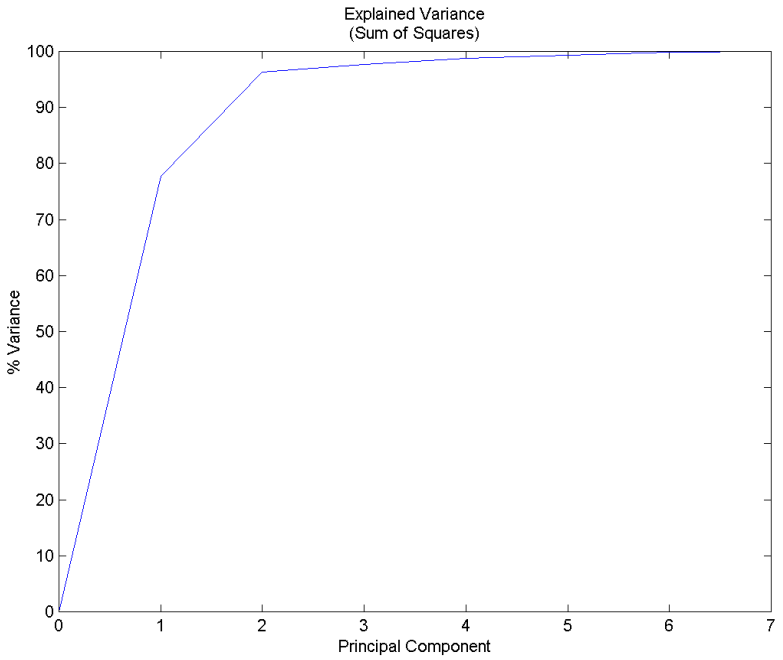


Figure 4. Variance modelled [%] per score-image.

The basic idea in MIA's score cross-plot set-up is to have one *comparison axis* with which to *interrelate* the whole series of cross-plots PC-12; PC-13, PC-14 etc. Adhering to such a standard set-up, one will need only a small number of cross-plots (one less the number of channels) in order to survey the gamut of all possible plots with e.g. PC-1 as the common X-axis. However there is also an additional number of higher-order cross-plots available of the type PC-23, PC24 ... PC-34, PC-35 .... through PC78 (in the present master data set case). It is important, especially for high(er) multi-channel work, not be unnecessarily bewildered by this potentially overwhelming array of additional cross-plots however. **In principle** everything there is to be known in score-space has been shown in the standard series with the one common anchor-axis. While it is indeed possible that certain higher-order cross-plots sometimes may serve to depict (very) specific pattern - in special cases - usually the inexperienced new user is greatly helped by first learning this systematic approach. Thus the standard cross-plot set-up includes only one higher-order plot, PC-23, Fig 2. It will serve the novel MIA user well specifically **not** to experiment with the higher-order option without some reflection and experience.

Colour-slice contouring in score cross plots - a 3-D histogram

All MIA score cross-plots employ a colour-slicing technique for depicting a three-dimensional histogram rendition of the relative number of pixels with identical score-pairs; for details see Esbensen and Geladi (1989) and Geladi and Grahn (1996). From the outside ("black sea surrounding the island"), in the score cross-plots the colour slicing grades olive/dark-green/green/yellow/orange/red/white, signifying that 0<5<15<45....>255 image pixels have been plotted at the same position in the pertinent PC-cross-plot, i.e. at identical score-pair coordinates in this plot. The exact numerical progression of the boundary values of these bins is actually only of minor interest; it's the overall visual impression of the *relative patterns and trends*, which carries the essential messages - very much in analogy to a *topographic map*, more of which below.

Brushing: relating different score cross-plots

The specific choice of the single cross-plot that is to serve as the starting point for a MIA analysis is very important. But what about the complementary cross-plots in which the same classes of pixels can also be displayed? *Brushing* comes to the fore. If no information to the contrary is present in a specific image analysis situation, it can be assumed as a working hypothesis, that the PC12 cross plot carries the most dominating (variance/covariance) information. This is so because these two first principal components carry the largest and second largest fraction of the total spectral space variance. In specific situations however, there is nothing against using any other, problem-specific combination of principal components images as the *starting score cross-plot*. In the present remote sensing example, if it was decided that we are specifically not interested in the overall reflectance aspect of the original image, this could easily be compensated for, simply by letting the analysis start out e.g. in the PC23 cross plot.

In this first illustration we shall make use of the standard PC12 cross-plot as the starting plot. Fig. 5 shows the technique of brushing, i.e. transferring a score space pixel class to the complementary other available score cross-plots. In this example we have delineated a rather large class "K". It will come as no surprise how the common PC-1 anchor axis in the PC13 and PC14 plots strongly guides the brushed dispositions of all pixels in the PC12 master class. One may perhaps appreciate the impression that in some of the plots the brushed class is "floating" above the main histogram. Note how the brushed class appears to "dilute" the complementary patterns in score space, as is quite possible, because the class was indeed defined in *another* score cross-plot.

This is a fairly typical result in a situation where one decides to start the analysis with one of the higher-order cross-plots, hopefully for a well-reflected reason (sic). With a little careful consideration the marked disposition in the PC23-plot might actually easily have been predicted directly from the PC13-plot.

With this standard powerful explorative brushing facility, it is possible to assess every *tentative* MIA-class in the gamut of all other potential score cross-plots - indeed one should always do so. Features observed only by brushing include "splitting" (one apparently coherent class, actually splitting up into two, or more, classes in higher-order

cross plots); "smearing" (obvious effect in the higher-order plots), "dilution" (illustrated above) i.a. We shall show several illustrations below.

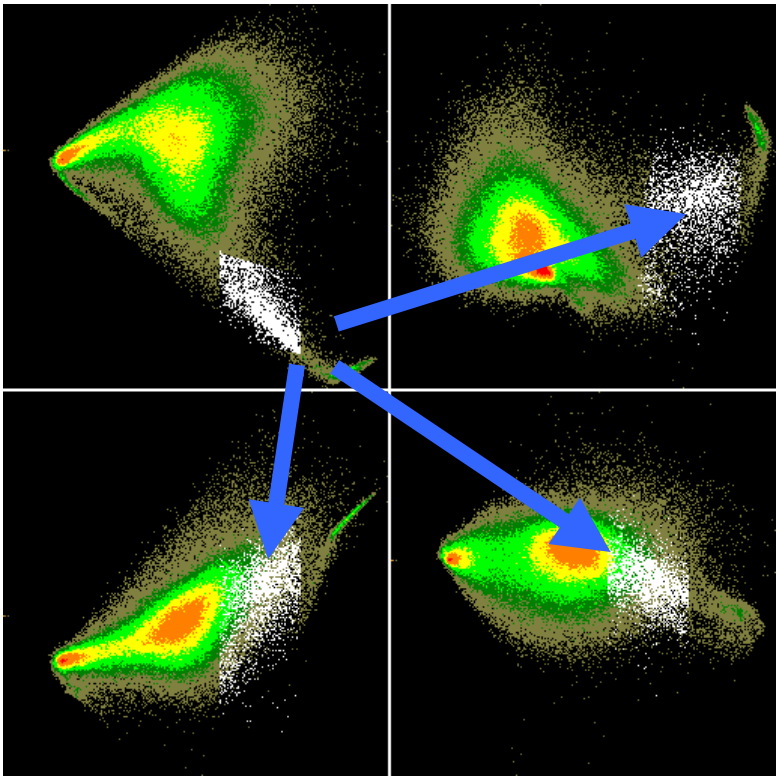


Figure 5. Brushing of MIA class "K", delineated in upper-left quadrant in score plot PC12

Joint normal distribution - or not

From extensive data analytical and statistical experience with principal component analysis, it is clear that only truly ellipsoidal pixel clusters in *all* PC-combination cross-plots can be said to meet the requirements of joint multi-normal distributions for all channels. It is thus very easy and uncomplicated to decide whether a particular class actually meets such requirements or not - and thus equally easy to find out whether unacceptable breaking of the premises of quite a number of traditionalist pattern recognition classifiers etc. will take place or not. Alas, we have yet to see many good examples of truly joint multi-normally distributed classes in nearly all types of multivariate imagery from science, technology and industry. True a very few have indeed been noted in our combined experience, but these cases are vastly overwhelmed by the many other types of strikingly non-normal distributions (multi-modal distributions), all of which can easily be analysed with the standard MIA approach however. MIA allows the user to make allowances for any specific class shape in the pertinent boundary delineations. Fig. 6 shows a relatively complex score aspect layout from this realm, which is a LANDSAT image from the Myvatn area in Iceland.

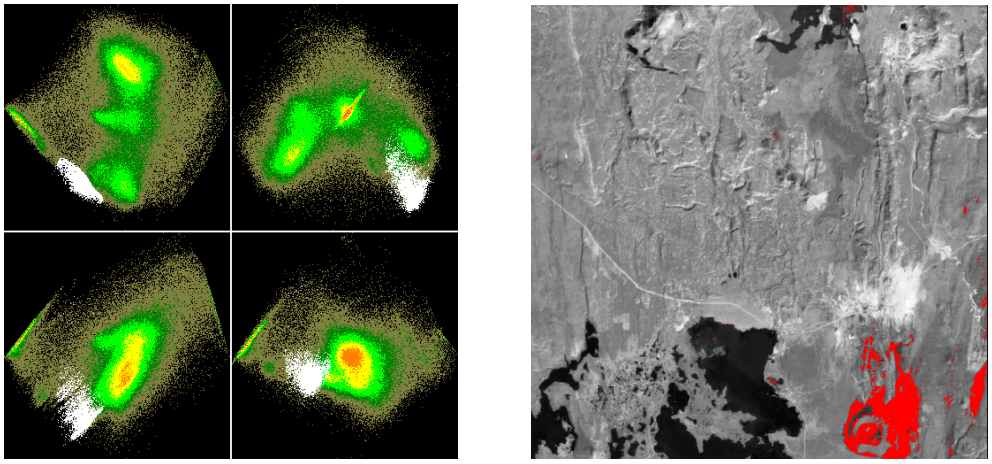


Figure 6. Fig. 6. Example(s) of complex, multi-modal layouts in score space. Original class delineated in the PC12-score plot (upper-left quadrant)

Local models/local modelling - the link to the classification *modus*

We are now in a position to introduce a major theme of MIA, that of so-called **local modelling**. The image analyst may for instance be interested in using just a subset of a scene, or of a score cross-plot aggregate, as a basis for a new, independent PC-model. Subsets are often on the agenda when it is not the entire square image upon which we wish carry out a new PC-decomposition in its own right. Geladi and Grahn (1996) develops the theme of local modelling in some detail. Reasons for such a local PC-modelling facility are invariably closely related to the specific image analytical objectives, which of course will vary from image to image.

Image subclasses come in two distinct varieties only:

1. Sub-classes delineated in scene space (traditionalist fashion), or
2. Sub-classes delineated in score space as *bona fide* MIA-classes

In this paper we have nothing more to say regarding the sub-optimality of the first category, which has been adequately denounced by Esbensen and Geladi (1989), Esbensen et al. (1993), Lowell and Esbensen (1993) and Geladi and Grahn (1996).

On the other hand, it is a very simple matter to direct MIA's PCA-module to work only on a selected score space class **as an alternative to the entire image**. Based on such a ***local model*** it's equally simple to follow up and let MIA calculate scores for all pixels in the image, said scores now corresponding to the covariance data structure of this local PC-model only (not corresponding to the entire image any more). This of course also applies to pixels in related scenes, images etc.

The concept of MIA local modelling is very useful for more advanced work, but proper understanding and competence is dependent upon a thorough understanding and experience of the basic MIA PC-modelling concepts first. Once this has been mastered however, there is really only very little difference working with *global* or *local* models.

The essential difference lies more with the specific reasons behind the need for local modelling. MIA's main contribution here again resides with its design primacy of delineating the appropriate local models in score space. Amongst other features, the local modelling feature can be shown to open up for the second-generation MIA concept of Multivariate Image Regression (**MIR**), which is a hot subject on the agenda for present R&D work at our laboratories, to be presented elsewhere (Lied, Geladi and Esbensen 2000). For a first introduction to the topic of MIR, see e.g. Geladi & Esbensen (1991), Esbensen et al. (1992).

Any local model, selected and delineated on the basis of a pertinent problem-specific reason, may serve as a basis for a re-classification of the entire image. It is important that completely new images (or relevant *parts* hereof) may now be classified, or re-classified as the need may be, in a completely analogous fashion to that of any global MIA model classification. This feature opens up for the complete range of discrimination/classification facilities of well-known features such as pattern recognition, SIMCA-classification etc. - which is an entire topic for itself and for a sequel paper.

MIA ANALYSIS CONCEPT – MASTER DATA SET ILLUSTRATIONS

The topographic analogy

In the following we shall adhere strictly to a topographic map terminology when discussing how to analyse the series of MIA score cross-plots. We shall use straightforward analogues: island, peninsula, peak, ridge, rise, flat, watershed, while also making use of *imaginary*, or virtual, "brooks, or rivers" supposed to follow *ditto* valley bottoms etc. We shall rely heavily on the reader's imagination in this endeavour, which is all-critical: The **topographic analogy** constitutes the core of the subject-matter of the central MIA principles exposition below.

The colour-sliced score cross-plots, Fig. 2, are specifically designed to be viewed, and interpreted, **exactly** like a topographic map. Thus e.g. white areas, which invariably will be situated only in the centre of the "topographic highs", signify the 3-D frequency *histogram peaks*, i.e. the highest densities of pixels with *similar score-pairs*. Uni-modal and multi-modal pixel distributions are revealed with absolute unambiguity. In Fig. 2, one thus observes three major topographic peaks (PC12 plot), more of which below. There is never any question about where and how these topographic peaks are to be found – it is not important that not all peaks boast a pixel density which necessarily results in white "snow-capped" peaks; it's dominantly the *relative* topographic expression which is important in this first interpretation stage (in fact there happens to be no "snow-capped" peaks in Fig. 2, but see later).

Much more important - more subtle pixel groupings and trends are also clearly outlined, never mind that such phenomena may be outlined in "only" the smallest of relative density terms, e.g. in the olive fringe areas only. A case in point in Fig. 7 is the very prominent south-east trending "ridge" in the lower half of the PC12 quadrant, termed "B". The relative proportion of all pixels encompassed by this ridge is actually far less than 2% of

the total number of pixels in the original image, while its *covariance trend* occupies a much more significant part of this plot. MIA is i.a. designed towards the greatest possible sensitivity w.r.t. this type of subtle features in the score space domain.

In fact, MIA's 3-D colour-sliced histogram score cross-plots often result in an "inverse" mapping of the frequency manifestation of the dominating scene data structures, such that these will be "compressed" into geometrically constricted "peaks" etc. in score space. Any class of significant *geometrical coherence* in the score cross-plots by necessity must represent a *bona fide* image pixel class, *irrespective* of the corresponding spatial/geometrical disposition, or apparent size in the scene. Examples abound in which there is very little, or no correspondence, between well-defined score space pixel classes and the archetype, spatially coherent *training classes* defined in scene space. Note that this situation is what prompted our critique of the "traditionalist image analysis" *scene space training paradigm* (Esbensen and Geladi 1989), Lowell and Esbensen 1993), Geladi and Grahn 1996), compared to which we here present the MIA alternative/complement.

A(ny) major histogram peak in the score-plot(s) necessarily corresponds to a (very) large proportion of image pixels, but it is often not very illuminating to focus MIA's attention on such prominent features - since they are simply manifestations of the (by far) most dominating image structures, which are always very clearly observed in the scene space anyway. A case in point is shown in Fig. 7, in which a MIA class of the absolutely most dominating peak in the Montmorency Forest scene, the central "volcanic peak" of the "volcanic island-like" PC12 score cross-plot analogy, termed "D", has been mapped back into the scene space. The forestry interpretation of this class is very clear: *undifferentiated* clearcut re-growth. The first general MIA rule: all dominating peaks in score space correspond to the dominating image space structures/segments. There is nothing new, nor even particularly interesting were this the only MIA feature. In fact MIA merges with "traditionalist" image analysts when these first-order, most dominating image structures/objects/segments are the only items on the agenda.

But this is also where MIA parts with the traditionalist image analysis concept, which - by definition - cannot delineate *subtle class features* by starting out in scene space, with anything even remotely akin to MIA's power, as shall be shown in full depth and detail immediately. In this situation MIA rather presents itself as a most powerful *complement* that specifically only claims rights of true *superiority* - and progressively so - when the more and more subtle details in the image comes to the fore. MIA comes on very strongly indeed for all weakly populated and/or subtly defined data structures in both scene as well as score space. Exploratory MIA image analysis is especially aimed at finding and highlighting exactly these types of subtle peak-structured or similar less well defined pixel aggregates (e.g. isolated "islands", "peninsulas", "ridges" i.a.), that otherwise run the risk of being *swamped* or *drowned* in the dominating structures and textures when delineated in scene space. Almost the remainder of this exposition is devoted to showing one or other aspect of exactly this, much more difficult-to-analyse image analysis situation, the *subtle class regimen.*

MIA will now be presented in a series of practical image analytical sessions. Along the way both an *exploratory image analysis mode* as well as a *pattern recognition (classification) mode* shall be illustrated, as shall demonstrations of other related image analysis objectives, which lends themselves naturally in the **MIA** context. The totality of

image analysis objectives/operations to be displayed below need not all be put into use simultaneously, nor always be all relevant for one particular image or scene. But we shall deliberately, and quite literally, take the Montmorency Forest scene apart in every which way in order to show to power of the general MIA approach.

MIA topographic score space delineation of single classes

MIA has been designed to allow the image analysts to focus on any "interesting" pixel cluster, prominent or subtle, by convex *polygons*, five of which are delineated in Fig. 7. Thus class "A" is the class encompassing all pixels with highest reflectance in the entire scene. It is emphasized that the specific delineation of pixel classes has been designed to allow for maximum freedom by the image analyst when outlining the enclosing perimeter of the (convex) polygons. Usually this type of convex polygon follows the topographic contours to a large extent, e.g. class "D". By way of contrast, class B has a very different geometric layout which suggest itself entirely by way of the covariance trend of the pixels involved however.

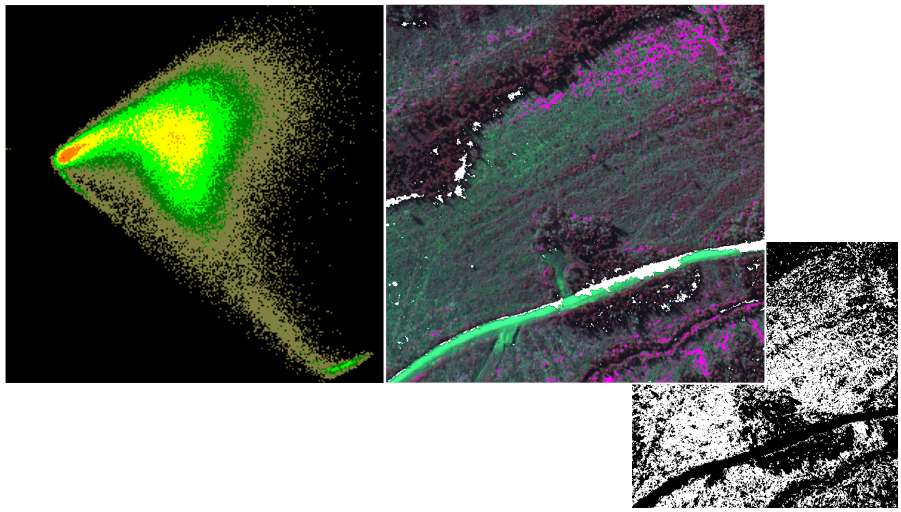


Figure 7. Maximum intensity/albedo contrast axis (Classes B --> A). Red pixels in scene space: "A". White pixels in scene space: "B". Insert: scene-space distribution of class "D" – undifferentiated re-growth .

All score space pixel classes are followed by immediate back-projection into the scene space. Either a simple binary mask is outlined with all designated pixels in white, as foe class "D", or usually MIA displays the original image *together* with the pertinent scene space overlay (in any suitable monitor combination R:G:B). Fig. 7 also shows the resulting spatial layout of this back-projection of the two pixel classes "A" and "B". Upon inspection, classes A vs. B turned out to represent the pixels in this particular scene with the absolute highest radiometric reflectance, class A (for all channels), and the opposite class of the absolute darkest pixels, class B. This latter class represent "shadow-pixels" - easily enough appreciated when the entire scene is viewed with particular notice of the general sun illumination direction (from NW), cmp. Fig 1. It is important to appreciate how the user is accorded complete freedom to *iterate* and *refine* this type of (tentative) class definition procedure as often as needed, should the first scene projection(s) indicate

only a sub-optimal class representation, as revealed in scene space. Domain-specific *interpretation* of the class masks in Fig. 7 may for example tell a forest expert who is familiar with the imagery and the general features of all the prevalent tree types in this particular context, that these resulting scene boundaries are not optimally delineated yet – in which case one more *iteration* of the score space class delineation is called for etc.

This score space/scene space *iteration* constitute the most important design principle of MIA in the explorative mode – an interactive interpretation stage, **from** score cross-plot pixel class delineation(s), **to** projected scene space class outlines, **complete with** original image underlay. It is up to the user, be it a domain specialist also versed in image analysis or a two-person team covering both these fields, to carry out this interactive procedure to as high an interpretation detail as deemed necessary by the image analysis objective(s). MIA's on-screen capabilities have been designed such that this interaction score-space/scene-space interpretation work can be as comprehensive and effective as needed, Lied (1999).

Working systematically, MIA analysis of this particular forest scene will soon reveal five primary classes in Figs 1 and 7, wiz. the dirt road, areas underlain by shadow, mature forest stands, undifferentiated re-growth, and a somewhat specific type of class of high albedo/reflectance.

Pixel class
A: areas composed by high reflectance pixels
B: areas in shadow
C: road (dirt road, not metalled)
D: regrowth (undifferentiated), mainly in the clearcut areas
E: Mature tree stands, also single crowns of old trees

These primary class designations will play a central role when MIA's next major image analytical features are to be developed, the *end-member mixing class* concept.

The classifications revealed in Fig 7 are mainly based on a more-or-less comparable equal-area basis in the score cross-plot. This type of grouping may however sometimes lead to misrepresentations, when it is remembered that equal areas in the score plot may in fact represent very large differences w.r.t. the actual number of pixels in the 3-D histogram bins, with density differences as large as 5:255, or more. Usually this "inverse" representation does not cause undue problems however, although it certainly pays to be aware of it.

The above illustrations stress the point that careful – *iterative* - pixel class perimeter delineation is of the outmost importance. We strongly believe that the topographic analogy is natural, especially as regards peaks etc. It is most likely the topographic expressions of "peakedness" that leads the human cognitive facility to form this type of *pattern cognition* very easily. But there is another, equally "natural" type of MIA-class now to be distinguished – an *end-member* series, or the mixing *class* series.

MIA delineation of end-member mixing classes

In Fig. 8 we have delineated three *mixing class series*. Note how the direction of these class delineations are directly related to the topographic ridge patterns (watersheds). In the case of class-series "X" and "Y" this terminology would appear obvious and very relevant.

E.g. classes "Y1-Y3" are comprised of the "mixing" ridge between classes mature tree stands "E" and the re-growth class "C" and likewise, class "X1-X6" can be seen as representing a mixing series between end-member classes "C" (re-growth) and the road class "D". We shall here also designate "Z1-Z3" a similar mixing class series, but now a mixing between the manifest end-members re-growth ("C") and a *virtual* high reflectance end-member "H". Observe how this similarity allows the image analyst to analyse *all types* of mixing-classes by relaying on only one common concept. Figs 9-10 shows two examples of the *gradual relationships* displayed by these mixing classes, especially when followed *from* one end-member *to* its opposite in scene-space.

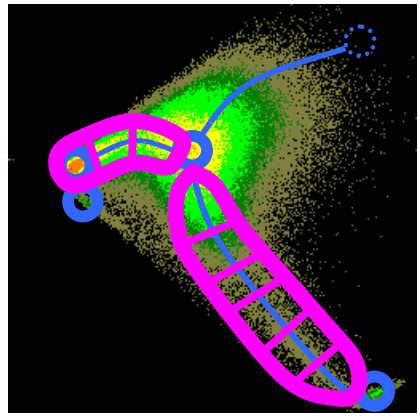


Figure 8. MIA mixing class concept (three mixing class series delineated). Each mixing class series extends between two appropriate end-members (some of which may be virtual)

In this mixing class delineation we have at first only placed emphasis on identifying the end-members making up the series extreme end points, but there are two additional much more penetrating and powerful interpretative "hidden information principles" behind these mixing-classes. Starting with the road class, following the "X" mixing series towards the central re-growth class and immediately *continuing* along mixing series "Y" ending up with the mature tree stand class – forestry interpretation of the entire spatio-temporal sequence leads to the following hypothesis:

The dirt road is regularly swept clean of all incipient re-growth saplings, because of the regular traffic of heavy duty forest machinery involved in the overall clear cutting operations a.o. – consequently the scene space pattern class for "X1", the first in the direction of the central re-growth class *per force* must represent whatever incipient re-growth can be observed in the scene. In other words, the first mixing sub-class immediately adjoining the road class must represent the *absolute youngest* re-growth saplings, with progressively older representatives forming the *grading sequence* of the scene-space rendition in the mixing class series X along the watershed route, ending up in the central re-growth class, "C". This can be easily appreciated in the scene-space rendition of Fig. 9.

By similar reasoning, the grading mixing class series Y can be *parsed* in a corresponding fashion– in an identical spatio-temporal context - i.e. **from** "C" **ending up** with the oldest, most mature trees standings in the scene, class "E", as is laid out in detail in Fig. 10. The longer one progresses along the Y-series (Y1-Y2-Y3), the older the trees delineated.

Observe how one is actually able to follow this entire *interpreted* growth process progress in minute detail in this score space rendition, Fig. 8, and immediately have access to its scene space dispositions, *vide* the sequenced imagery in Figs 9, 10.

Thus the entire *ensemble* of score space classes $X_1 \rightarrow X_6 \rightarrow C \rightarrow Y_1 \rightarrow Y_3$ represent a *spatio-temporal slicing* of what could be interpreted (and termed) the "re-growth process" in the context of this particular scene. It is essential to appreciate that this interpretation takes place by starting out in score space, but it is only when the resulting MIA-classes are displayed in scene space, that full interpretation of their meaning is possible. Also this interpretation is validated mostly by reference to other, already interpreted or well-segmented features in the scene. In this particular dynamic, multi-temporally affected scene, the biologic process of *re-growth* has be subjected to a kind of "stroboscopic" time-slicing, delineated by the sequencing of juxtaposed mixing-classes.

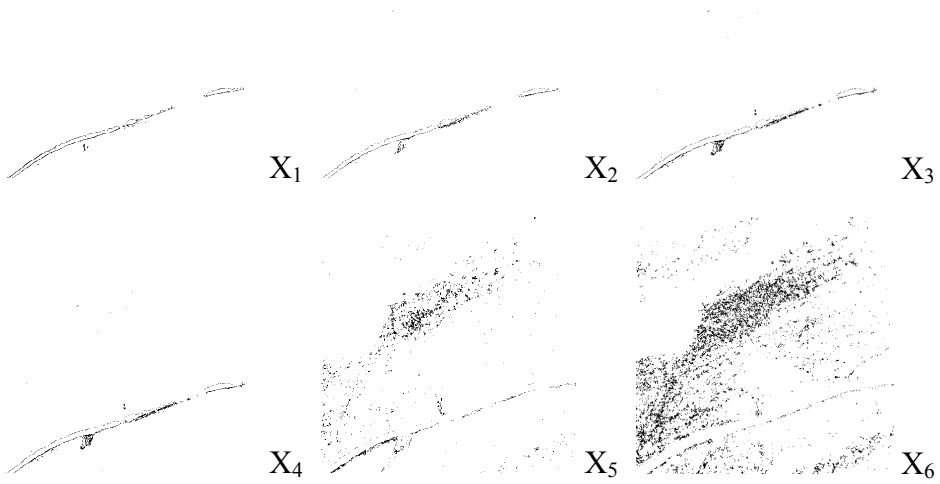


Figure 9. Scene space disposition of mixing classes X1 - X6.

By this development of the end-member/mixing class concept it has been possible to shed hitherto unimagined detailed light on this complex spatio-temporal re-growth process. Clearly it is the *informed interpretation* of the domain specialist – and full command of MIA's capabilities – which underlies this powerful *analysis*. MIA allows a temporal-spatial decomposition which is unparalleled in traditionalist image analysis; there is simply no possibility to decompose the multivariate image in similar segments, were this to start out from the scene space, in which the would-be "training classes" are hopelessly far too *disjunct* and far too scantily distributed, as dramatically laid in their scene space context, Figs 9, 10.

There is one more, phenomenologically identical, mixing series present in the score space rendition of this scene, but with a distinctly different non-biologic interpretation, the $Z_1 \rightarrow Z_3$ mixing-series, situated in an almost "perpendicular" disposition w.r.t. to the re-growth series. This series represents the ultimate span of the general low→high intensity

(comp. the above "shadow-reflectance" contrast phenomenon, Fig. 7), but in the present context it can be seen as also tracking *across* the same central re-growth class. The *physical interpretation* of this new axis remains the same: the "$Z_1 \rightarrow Z_3$ vector" must represent a *generic*, presumably mostly physical reflectance direction (in score space), signifying gradations in the total reflectance recorded from within the confines of one ground trace pixel size (70 cm x 70 cm). The reflectance in this type of imagery is surely dependent on a composite set of factors, among which individual leaves, their colour, angle w.r.t. the sun illumination, degree of moisture coverage i.a. plays important roles.

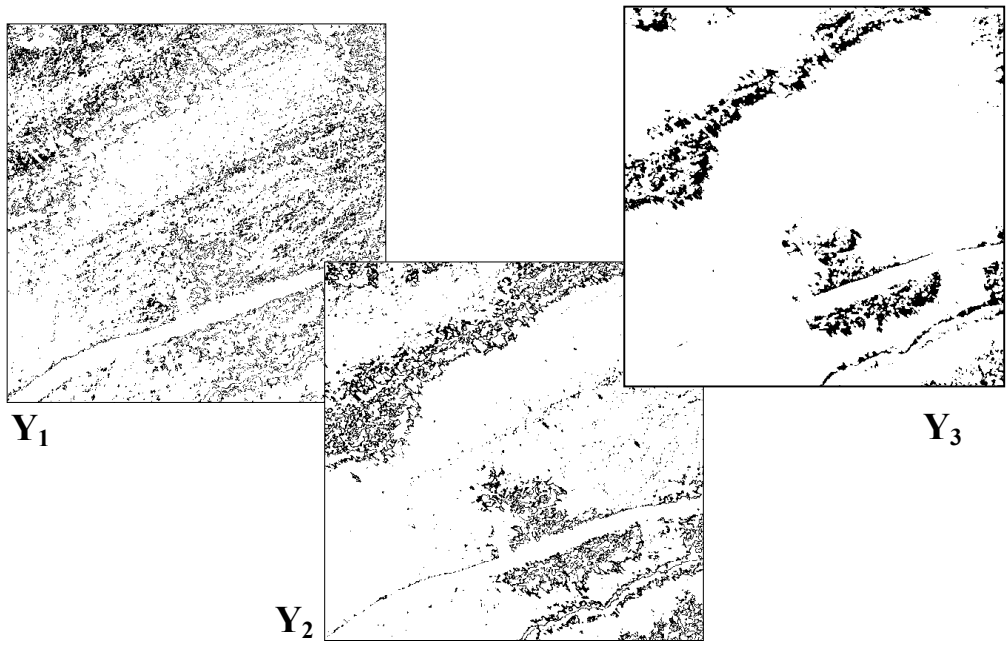


Figure 10. Scene space disposition of mixing classes Y1 - X3.

There are thus three general aspects of this detailed MIA analysis to be highlighted.

1. Observe how these two distinctly different compound mixing-series lie very close to "orthogonal" to each other. This is of course no coincidence, but is simply an inheritance from the underlying PCA analysis, the design purpose of which precisely is to decompose covariance trends according to forced *orthogonal axes* (principal components). In this particular case it is mainly PC1 which is rightfully interpreted as the dominating general intensity (albedo) axis, though the slightly oblique direction of the "shadow-reflectance" axis ($Z_1 \rightarrow Z_3$) bears witness to a slight involvement of PC2 as well in delineating the direction of the overall increasing albedo.

2. Observe how both these genetically different series meet, or cross over, at the scene's singular most dominant class, the central re-growth class "C". This hammers home the message that class "C" is nothing but a veritable "mixed bag" of many "types" of re-growth manifestations and this also explains why the simplest MIA-class analysis of delineating the class C peak in this case results in an overwhelming number of scene space pixels, drowning out most, if not all, possibilities of making detailed interpretations

of the nature of this class - other than that of the standard forestry "undifferentiated re-growth" category. In fact this type of cross-over class, this type of central, dominating class is very often met with in MIA analysis of imagery, certainly not only from (high-resolution) remote sensing, but from many other technological and industrial imagery as well. MIA constitutes the *only* image analysis tool with sufficient power for these kinds of complexities (sic).

3. We have now presented the two, partly alternative, partly overlapping *meta-principles* for MIA pixel class delineations: individual "peaks" vs. "mixing classes".

Which to use? – When? – How?

One should not dismiss the above "biologic growth interpretation mode" of analysis to quickly even if this type of dynamic time-slicing may at first sight seem rather specific for remote sensing in general, for change-detection in particular. Very probably however, this type of "change-analysis" constitute a *generic type of interpretation*, which can be modified and applied in several other image type contexts as well. Certainly this is the experience of the authors from a suite of very different technological and industry-related types of imagery. In any event, the other "non-dynamic" mixing series more reflects a static, scene-dependent physical phenomenon (albedo/intensity contrast/reflectance), the like(s) of which will be present in almost any multivariate image from science, technology or industry in which (natural or artificial) illumination plays a role.

In situations where this mixing-class concept would not appear to be useful, a return to the simple "peak" delineation of what will surely always be representative, objective pixel classes will often be quite sufficient then. The perhaps most interesting field here would be the *interplay* between these two types of class delineation principles, but needless to say, this problem will always be "scene-specific" to a very high degree. One must never forget that each new multivariate image always should be analysed on its own accord. However, experience with some several *types* of multivariate images from a vast range of different origins (remote sensing imagery, several types of technological imagery, industrial...) reveals that the gamut is indeed made up of peaks and end-member/mixing-classes, to more than 95%. - So when to use which approach: peaks vs. mixing series?

The resolution lies in the fact that all mixing-class series should be subjected to identical mixing-class analysis as that presented above, and that any end-member always also constitutes a legitimate single peak in its own right. By employing the concept of end-member mixing/classes (directed along the watershed ridges, connecting peaks in the topographic analogy setting) all types of connected peaks are in principle always open to either type of analysis. But clearly it is the *scene space knowledge* that will determine whether it will be possible to make *meaningful interpretations* of the "mixing-sliced" subclasses.

In this paper we have delineated two generic types of mixing classes. We would probably be grossly simplistic, were we to suggest that this is all there is; that still other "types" of analogous analysis axes will not be found to be associated with other type(s) of imagery in future applications, but we have actually not yet found the need for additional image analysis concepts. We are by now fairly certain that the suggested concepts, or meta-

principles, of end-member/mixing class vs. standard "peak" MIA analysis will be of very general, perhaps even universal, applicability.

Scene-space sampling in score space – a final detail

Often the image analyst may need to sub-sample an entire image in the specific meaning of a *representative sub-sample* of the image. This may be on the agenda for many different reasons, e.g. for forest inventory purposes, where a forester would like specifically to sample *all known* forest and growth *classes* present. One specifically needs to be certain that all classification classes indeed are *equally* represented on a spatial basis. Many procedures and sampling schemes have been developed over the years for this and related purposes, all of which operate in the scene space in the image, or on the map. There are many parallels to this sub-sampling situation from other types of imagery as well.

Interestingly MIA may also here offer an alternative to this scene space tradition. Again we illustrate using the clearcut study imagery. In Fig. 11 observe the extremely thin class delineated in the lower left quadrant - actually this class is only one pixel wide. This class is immediately brushed into the three complementary PC cross-plots. It is of significance that we here have made use of the PC23 cross-plot for the class delineating purpose, for reasons that will become immediately clear. Note first how this PC23-class covers *all* the major classes present in score space - here we actually take advantage of the fact that most of the major classes are *non-resolved* in the higher-order, e.g. PC23-plots. It is precisely because of this judicious use of the PC-23 cross plot that we have been able to acquire a representative, complete, equal-density sampling of the entire covariance data structure in feature space by using the simplest of class delineations – a line, as is indeed substantiated in the accompanying PC12, PC13 and PC14 cross-plots. It is especially gratifying to observe the inherent "splitting" in the PC13 (upper right) plot. This one-pixel thin sampling class has done a remarkable job sampling over all classes indeed!

Fig. 11 also shows the corresponding spatial projection. Indeed a uniform spatial disposition of potential inventory localisations has been achieved, complete with a number of denser structures present (major tree stands etc.). It is now a simple matter to overlay this display with e.g. a road map or the like and to proceed with a logistical planning for the forest inventory, in which a further weeding out of "surplus" sampling sites no doubt will form an integral part. It's really not a problem worth mentioning weeding down an already acceptable spatial template - relative to the opposite case. Thus, for such sampling purposes one might advantageously seek out the *least structured* score cross-plot.

There are other variations on this sampling feature of MIA, all invariably related to the specific image analysis problems at hand and their special objectives, but we leave it to the reader to associate freely from this generic example.

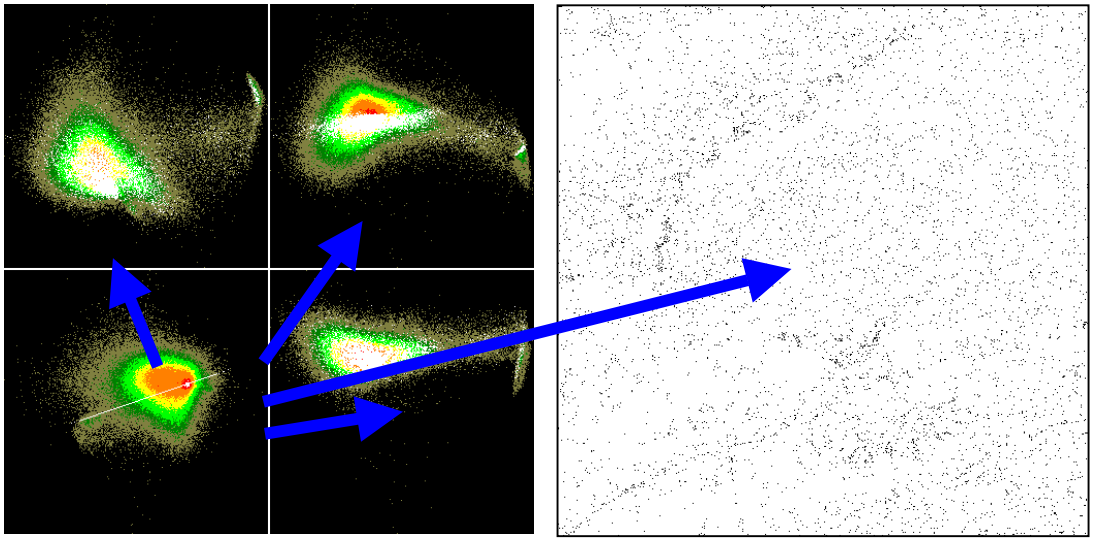


Figure 11. Generic MIA sub-sampling concept. From score space to scene space: equal density sampling w.r.t. the image covariance structure(s), not w.r.t. scene-space area density.

CONCLUSION

Multivariate Image Analysis (**MIA**) is a set of *interdependent image analysis procedures*, encompassing both an explorative and pattern recognition mode, as well as classification-related facilities. A synoptic overview of all elements in MIA has been presented by a series of different applications on a comparison data set. We presented extensive justifications for the specific MIA approach - *contra* the traditionalist image processing mode.

MIA is a set of problem-dependent, *interdependent* explorative analysis and classification tools to be used by the *informed* analyst, according to the *specific image analysis objectives*. This review furnished a generic overview of the *principles of MIA* needed to compose one's own flexible, problem-specific strategy of multivariate image analysis: Mandatory pixel class delineation in feature space (*score space*) by a topographic analogy and a dual end-member/mixing class vs. peak class delineation concept, of *universal applicability* also in other image modes.

MIA constitutes a most powerful image analytical concept for dealing with *any degree* of complex imagery: MIA's design – and analysis principles are *invariant* w.r.t the number of channels present. With MIA is not necessary to invoke massive, parallel computer approaches in order to deal with even the most complex imagery (Lied, Geladi & Esbensen 2000).

REFERENCES

MCCOLL, W.D., TILL, S.M. & NEVILLE, R.A. 1984: MEIS II operational sensor for multidisciplinary studies. 9th Canadian Symposium on Remote Sensing, pp. 497-501.

ESBENSEN, K.H. & GELADI, P. 1989: Strategy of Multivariate Image Analysis (MIA). Chemometrics and Intelligent Laboratory Systems, vol. 7, pp. 67-86.

ESBENSEN, K, GELADI, P. & H. GRAHN 1992: Strategies for multivariate image regression (MIR). Chemometrics and Intelligent Laboratory Systems vol. 14, pp. 67-86.

ESBENSEN, K.H., EDWARDS, G. & ELDRIDGE, N.R. 1993: Multivariate Image Analysis in forestry applications involving high resolution airborne imagery. 8.th Scandinavian Conference on Image Analysis - SCIA'93, pp.953-963.

LOWELL, K. & ESBENSEN, K.H. 1993: Is image segmentation really a valid technique for ... 8.th Scandinavian Conference on Image Analysis - SCIA'93, pp. 973-979.

KRAMER H.J. 1996: Observation of the Earth and its environment. Survey of missions and sensors (3.ed). Springer-Verlag. ISBN 3-540-60933-4

GELADI, P. & K. ESBENSEN 1991: Regression on multivariate images: principal components regression for modelling, prediction and visual diagnostic tools. Journal of Chemometrics , vol.5, pp. 97-111.

GELADI, P. & GRAHN H. 1996: Multivariate Image Analysis. (Chichester: John Wiley & Sons)
ISBN 0 47193001 6

LIED, T. 1999: MIA software. http://www-pors.hit.no/tf/forskning/kjemomet/kjemomet.html

LIED, T., GELADI P. & ESBENSEN K. 2000: Multivariate Image Regression, Implementation of Image PLS – First Forays. Journal of Chemometrics, 14 (2000) 585-598.

Paper II

JOURNAL OF CHEMOMETRICS
J. Chemometrics 2000; **14**: 585–598

Multivariate image regression (MIR): implementation of image PLSR—first forays

Thorbjørn T. Lied[1]*, Paul Geladi[2] and Kim H. Esbensen[1]

[1]*Telemark University College, Porsgrunn, Norway*
[2]*Umeå University, Umeå, Sweden*

SUMMARY

In the effort of analysing multivariate images, image PLS has been considered interesting. In this paper, image PLS (MIR) is compared with image PCA (MIA) by studying a comparison data set. While MIA has been commercially available for some time, image PLS has not. The kernel PLS algorithm of Lindgren has been implemented in a development environment which is a combination of G (LabVIEW) and MATLAB. In this presentation the power of this environment, as well as an early example in image regression, will be demonstrated. With kernel PLS, all PLS vectors (eigenvectors and eigenvalues) can be calculated from the joint variance–covariance ($\mathbf{X}'\mathbf{Y}$ and $\mathbf{Y}'\mathbf{X}$) and association ($\mathbf{Y}'\mathbf{Y}$ and $\mathbf{X}'\mathbf{X}$) matrices. The dimensions of the kernel matrices $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$ and $\mathbf{Y}'\mathbf{X}\mathbf{X}'\mathbf{Y}$ are $K \times K$ (K is the number of **X**-variables) and $M \times M$ (M is the number of **Y**-variables) respectively. Hence their size is dependent only on the number of **X** and **Y**-variables and not on the number of observations (pixels), which is crucial in image analysis. The choice of LabVIEW as development platform has been based on our experience of a very short implementation time combined with user-friendly interface possibilities. Integrating LabVIEW with MATLAB has speeded up the decomposition calculations, which otherwise are slow. Also, algorithms for matrix calculations are easier to formulate in MATLAB than in LabVIEW. Applying this algorithm on a representative test image which shows many of the typical features found in technical imagery, we have shown that image PLS (MIR) decomposes the data differently than image PCA (MIA), in accordance with chemometric experience from ordinary two-way matrices. In the present example the **Y**-reference texture-related image used turned out to be able to force a rather significant 'tilting' compared with an 'ordinary MIA' of the primary structures in the original, spectral R/G image. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: multivariate image regression; MIR implementations; multivariate image analysis; MIA; kernel PLS

INTRODUCTION

Since the introduction of multivariate image analysis (MIA) in 1989 [1], multivariate image regression (MIR) has not been developed to the extent one would have perhaps expected. The reasons for this might be low interest within scientific society, few inspiring MIR applications and/or lack of the required computing power. With the presentation of kernel PLS, however, Lindgren [2] has shown that it is possible to reduce this last factor significantly. Computing the PLS loadings using

* Correspondence to: T. T. Lied, Telemark University College, Kjølnes Ring 56, N-3914 Porsgrunn, Norway.
E-mail: thorbjorn@lied.no

Received 15 November 1999
Accepted 3 April 2000

only small covariance matrices instead of large multivariate images reduces the number of calculations tremendously.

In traditional two-way multivariate image analysis each pixel is looked upon as an object. In image analysis the number of pixels (N) is often large, and as technology develops, constantly increasing. Thus having e.g. two million objects is not unusual today. The number of variables (K), e.g. image channels, is usually very much lower, representing e.g. wavelength (colour), polarizing angle or frequency. When these types of multivariate images are unfolded [3], we tend to get very long and narrow matrices. In MIA the loadings are usually calculated using SVD (singular value decomposition) on the covariance matrix $\mathbf{X}'\mathbf{X}$ [4], which is a $K \times K$ matrix. In kernel PLS the loadings are calculated from the $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$ matrix, which is also a $K \times K$ matrix. Using only small matrices in the updating of this kernel means that one does not have to carry around the large $\mathbf{X}$ and $\mathbf{Y}$ and long latent variable vectors in the numerical calculations.

MIA is first of all intended for explorative image analysis purposes. Transforming multivariate images to their most important structures (latent variables) enables a dynamic segmentation approach with problem-dependent interpretation of similar objects in the entire image [1,4]. However, in situations where external knowledge ($\mathbf{Y}$-image) is available, image PLSR can now also be considered, based on its power in guiding the decomposition of the multivariate $\mathbf{X}$-image. For predictive purposes the use of some kind of regression model is required. Some very meaningful candidates are PCR [5] and PLSR [5–8]. In this paper an implementation of multivariate image PLSR, some considerations of the method and an early application example are presented. Other application examples are available [9]. Comparison of detailed results from PCR and PLSR will be presented in a future paper.

METHOD

Traditional algorithms [10] for calculating PLS scores and loading weights for a given PC carry around the large $\mathbf{X}$ and $\mathbf{Y}$ residual matrices and corresponding parameter vectors. Because multivariate images consist of very large matrices, typically two million pixels by K variables plus one or more $\mathbf{Y}$-variable(s), these algorithms consume enormous amounts of computer memory and processing time. Thus a different approach is desired for multivariate image data.

In 1994, Lindgren [2] introduced a method designed to reduce the matrix sizes during calculation. This method initially calculates three small kernel matrices, $\mathbf{X}'\mathbf{X}$, $\mathbf{X}'\mathbf{Y}$ and $\mathbf{Y}'\mathbf{Y}$, and the master kernel $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$. Loadings and weights are calculated using the master kernel, which in turn is updated for each component calculated, using $\mathbf{X}'\mathbf{X}$ and $\mathbf{X}'\mathbf{Y}$. Compared with the traditional approach, which needs to update the large $\mathbf{X}$ and $\mathbf{Y}$ residual matrices, the kernel algorithm can save tremendous amounts of memory, as illustrated in Figure 1.

This approach is based on the fact that scores and loadings can be calculated as eigenvectors using square kernel matrices:

$$\mathbf{w}\lambda_1 = (\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X})\mathbf{w}, \quad \mathbf{w} : \text{PLS } \mathbf{X}\text{-weights}$$

$$\mathbf{q}\lambda_2 = (\mathbf{Y}'\mathbf{X}\mathbf{X}'\mathbf{Y})\mathbf{q}, \quad \mathbf{q} : \text{PLS } \mathbf{Y}\text{-weights}$$

$$\mathbf{t}\lambda_3 = (\mathbf{X}\mathbf{X}'\mathbf{Y}\mathbf{Y}')\mathbf{t}, \quad \mathbf{t} : \text{PLS } \mathbf{X}\text{-scores}$$

$$\mathbf{u}\lambda_4 = (\mathbf{Y}\mathbf{Y}'\mathbf{X}\mathbf{X}')\mathbf{u}, \quad \mathbf{u} : \text{PLS } \mathbf{Y}\text{-scores}$$

Because MIA and MIR operate on vectorized images where $N \gg K$, $\mathbf{w}$ is a preferred starting point in the calibration procedure. In situations where $K \gg N$, this is not the case, because $\mathbf{X}'\mathbf{Y}\mathbf{Y}'\mathbf{X}$ becomes

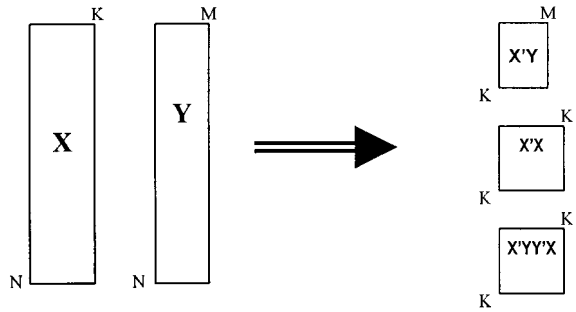


Figure 1. The kernel approach saves lots of computer memory required for calculating the weights and loadings. The actual amount saved is dependent on the *N/K* ratio.

very large. Instead, $\mathbf{XX}'\mathbf{YY}'$ is used [11] for this purpose. In situations where $N \approx K$, kernel PLS does not give much improvement. In Reference [12], kernel PLS is compared with a similar algorithm for the singular value decomposition of $\mathbf{X}'\mathbf{Y}$.

IMPLEMENTATION

It was found convenient to use LabVIEW as a programming environment for MIA/MIR. LabVIEW is mainly used for user interactions and file management, while MATLAB takes care of the actual number crunching. Our choice was made based on prior knowledge of LabVIEW and MATLAB as cost-efficient with regard to development time. The price we have to pay is a slightly slower algorithm than would be possible to obtain using C/C++. Especially the link between LabVIEW and MATLAB is slow when passing large matrices. The speed obtained is quite adequate for R&D as well as routine MIR, however.

LabVIEW (National Instruments website: www.ni.com/labview) is a graphical programming environment, written in C, which in the last few years has gained popularity and usability in numerous fields of applications. As the environment itself is becoming more stable and debugged, different toolboxes pop up around the world, introducing more and more pre-programmed functions, or VIs (virtual instruments) as they are called in LabVIEW. Because LabVIEW uses a graphical

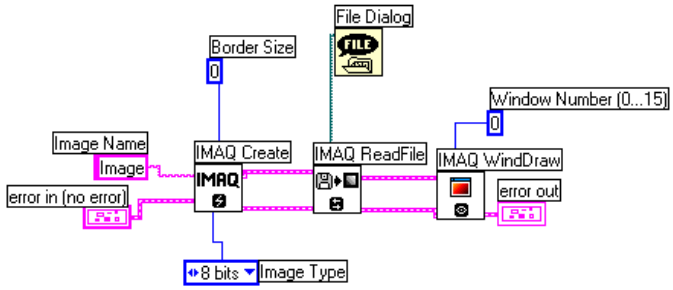


Figure 2. LabVIEW uses graphical symbols for different functions and sub-VIs, and the programmer connects these together using wires. User controls and indicators also show up as symbols in the diagram.

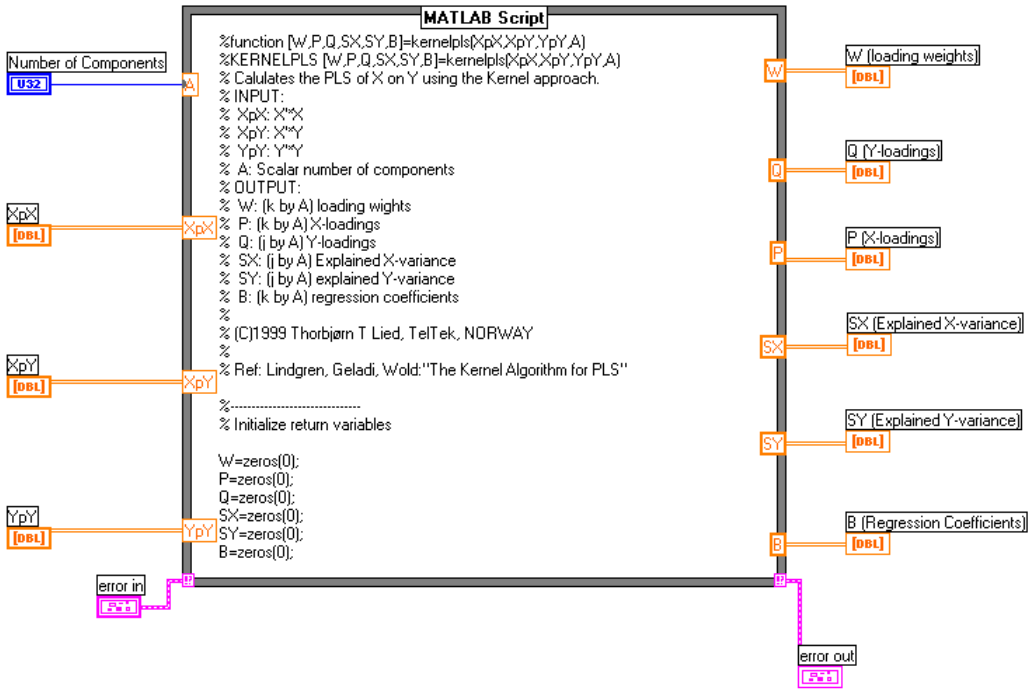


Figure 3. The actual diagram for the kernel PLS implementation. Note that the entire script is not shown as it is inside a scrollable box. The algorithm is found in Reference [2].

programming language, called G, and user interfaces are drawn directly in panels, LabVIEW truly is a visual programming system (Figure 2).

MATLAB (MathWorks Inc. website: www.mathworks.com) was used for the core numerical calculations.

There are two possibilities when combining LabVIEW and MATLAB. One is to put MATLAB scripts directly in the LabVIEW diagram, the other is to call external scripts (m-files) from the diagram. The first alternative was chosen here.

To reduce the amount of data passed between LabVIEW and MATLAB, it was decided to calculate the initial kernels in LabVIEW and pass these to MATLAB, which in turn returns loadings, loading weights and regression components.

As in most modern programming environments, in LabVIEW it is desirable to build each program as a collection of reusable sub-programs, or sub-VIs. This makes the code, or diagrams, easier to read and debug, and is of course indispensable for building complete dedicated software packages. Figure 3 shows how LabVIEW passes the initial kernels to MATLAB and calls upon MATLAB to perform the PLS calculation.

One level higher in the program, this VI is called upon with the kernel matrices as parameters. The implementation of this is shown in Figure 4.

Prior to the VI shown in Figure 4, scaling and centring of **X** and **Y** can be applied, if necessary. Following this VI, **X**- and **Y**-scores are calculated by projecting **X** and **Y** on their corresponding loadings. After this, we are ready to display score plots and score images, as well as loading plots,

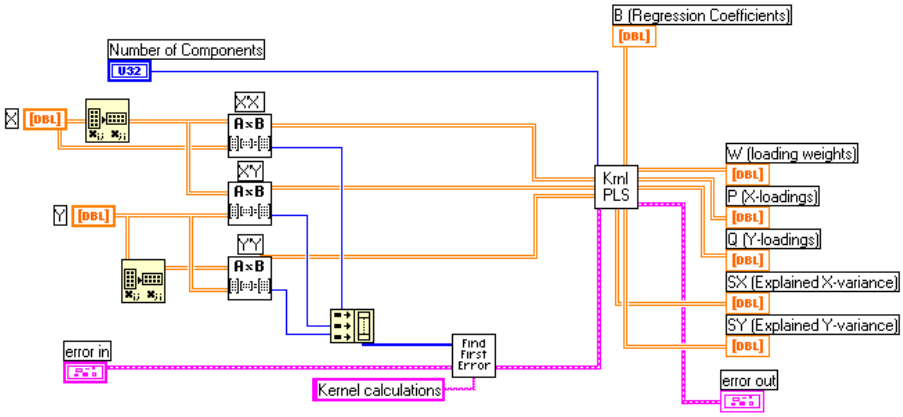


Figure 4. How the kernels are calculated and passed to the kernel PLS VI. The current VI is typically called upon after scaling and/or centring of **X** and **Y**.

and/or to carry out further calculations, etc., very much in the same tradition as with MIA [1,4–8].

While score images are shown in their original size and geometry, score plots are not. A score plot is a 2D histogram, or a scatter plot between two score vectors. Treating each pixel in the scene space as an object, and thus plotting each object individually, this scatter plot becomes unreadable without using an intensity colour-slicing map [1]. Score plots are used for object classification (**X**-scores, **T**), while **T** vs **U** (**Y**-scores) plots are used for evaluating the prediction performance of the image regression model [5–8]. The difference between scatter plots and images will be more evident when looking at an example.

APPLICATIONS

Because most effort has been put into the present new software development, a large range of applications is not yet available. Nevertheless, an early example with three spectral channels in the raw image is presented here. Although this is only a very modest multivariate image, it serves the purpose of showing the principles of MIR completely.

The raw image was captured with SILVACAM (VTT Automation website: http://www2.vtt.fi:82/aut/rs/prod/silvacam.html), which is a modified RGB video camera where the blue channel has been replaced with an NIR (near-infrared) channel. The composite raw image (R/G/NIR) is shown in Plate 1. In the present example, however, the NIR channel did not contribute much to the decomposition and was therefore removed from the data matrix for the texture derivations to be presented below (in order to give more room for the latter).

This image has been specifically designed to highlight both spectral as well as different textural and structural differences between the different objects in the image. Thus we have constructed an image with only three principal objects present:

- highly textured cloth as background (Canadian lumber-jacket);
- flat plastic fragment (‘prison window bars’);
- eight lead pencils in four colours.

The idea behind this image is that there are important differences between the spectral objects (which

can be discriminated by a standard MIA spectral decomposition [1,4]) and the textural objects which will be the main focus of concern in this application example (texturally there is e.g. only one type of pencil, while there are four spectral classes corresponding to the four colours).

Observe how the green image apparently conveys more detail and focus than the red image, especially regarding the definition of the highly textured background (Figure 5). The image is also representative of various forms of specular reflectance. This latter is directly dependent upon illumination angles, etc. For this constructed image a partly asymmetrical illumination was used, producing a clear light/shadow contrast primarily in the N–S direction.

Thus, while very simple in the number of object types present, this image in fact catches many of the principal image analysis elements and features of technological imagery, a number of different spectral classes, many or all with individual texture, illumination (light/dark/shadow) differences, etc.

The goal of this example is twofold:

(1) to discriminate between these different types of classes specifically with help from the textural information;
(2) to compare MIA vs MIR.

In order to do this, a new MIX (multivariate image texture analysis) concept is introduced whereby a series of textural image derivatives is directly added onto the series of spectral variables (from the perspective of both MIA and MIR, this simply results in a set of added **X**-channels). This will be done in three different ways in the present case.

Thus, for each of the two spectral channels (red and green), three relevant textural derivatives have been calculated, giving a total of $K = 2 \times (1 + 3) = 8$ channels (see Figure 5). The following texture filters were applied:

- median filter;
- Laplace filter;
- compound filter (sculpture + variance + median + inversion).

A reference **Y**-image is of course required for image PLSR. A 'texture index' **Y**-image (**TI**) is devised (Figure 6) which expresses the basic texture differences between the three texture classes in a quantitative manner. Texturally the piece of plastic is almost completely 'flat' (**TI** = 0–10); the pencils are slightly more complex texturally speaking (octagonal cross-section), resulting in **TI** = 20–40; while the highly textured Canadian lumber-jacket cloth displays a very high texture index, **TI** = 225–255. Figure 6 shows these texture relationships very clearly. This is the type of information that will be used in order to introduce textural relationships in the image decompositions, but exclusively as **Y**-information.

The **TI** image was constructed in Image Pro Plus from Media Cybernetics, applying a combination of texture-sensitive filters to the red channel in **X**. The combination consisted of 'sculpt', 'Sobel', '5 × median 5 × 5' and contrast enhancement, which, when applied in the mentioned order, gave the result shown in Figure 6.

Application MIA vs MIR—objectives

In order to see how image PLSR (MIR) decomposes differently than image PCA (MIA), three cases will be studied, in which the PCA and PLSR algorithms will be applied essentially to the same data set but in three different ways:

- case 1—MIA_0 (without **Y**-reference in **X**);
- case 2—MIA_Y (**Y**-reference included in **X**);
- case 3—MIR (**Y**-reference used in **Y**-block).

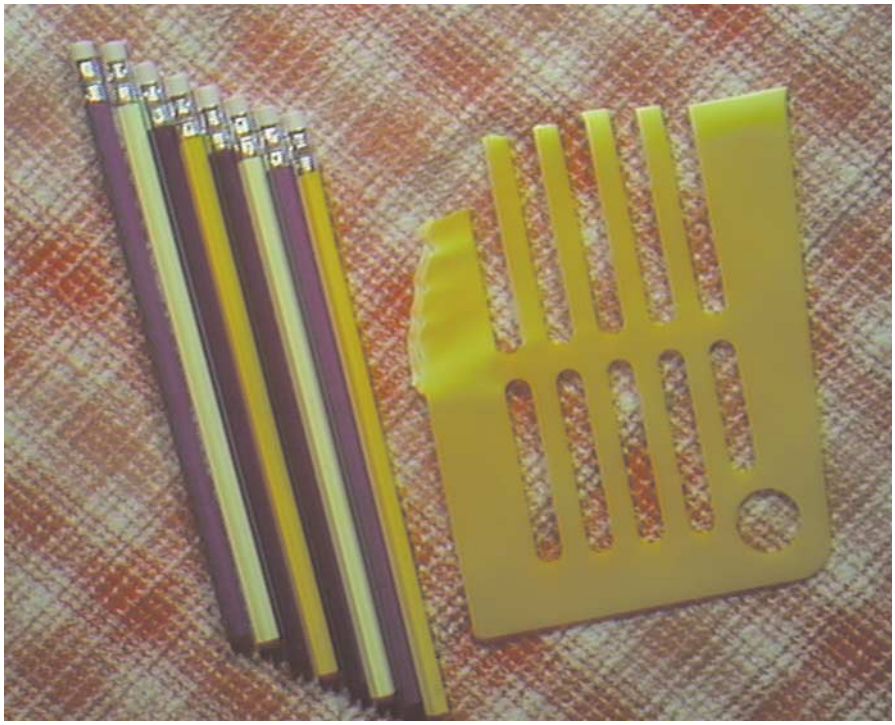


Plate 1. Raw image in three spectral channels, Red, Green, and NIR.

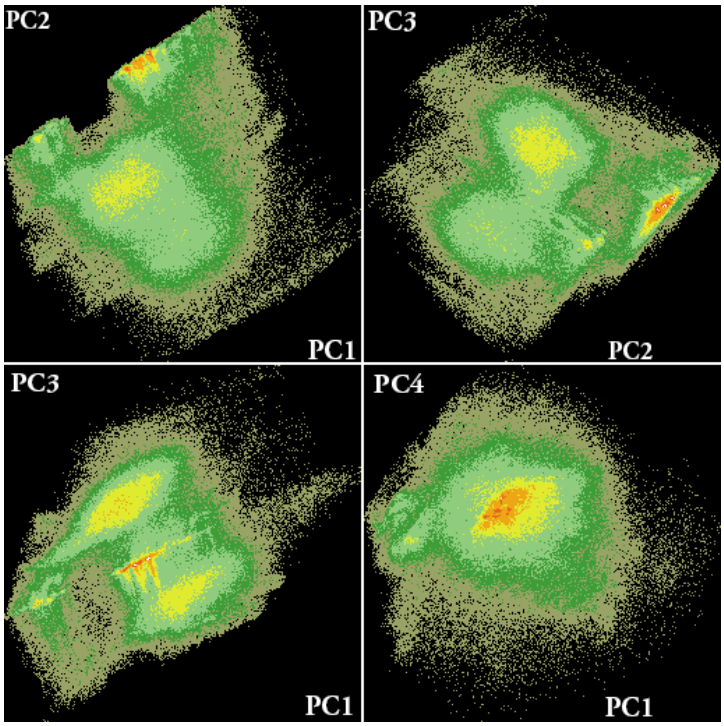


Plate 2. MIA_0 Scoreplots: 1-2, 2-3, 1-3 and 1-4.

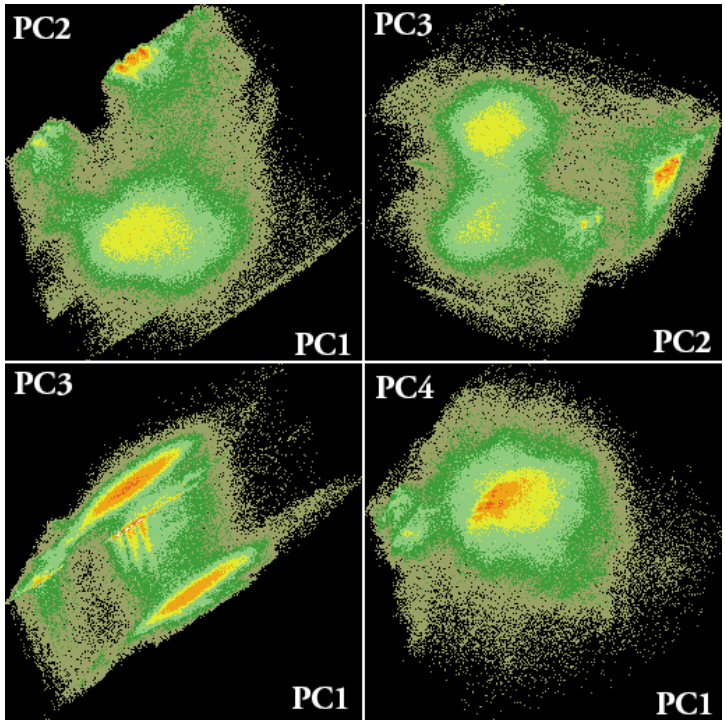


Plate 3. MIA_y Scoreplots: 1-2, 2-3, 1-3 and 1-4.

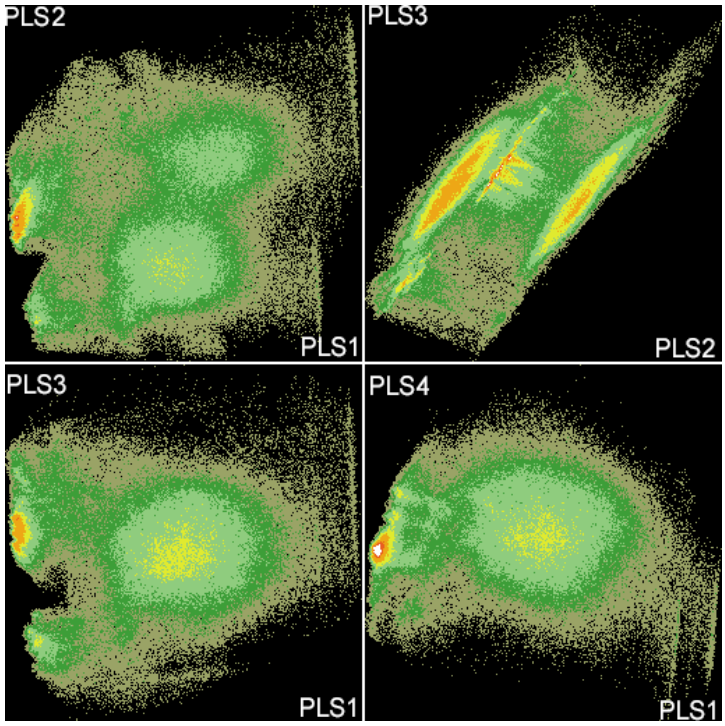


Plate 4. MIR Scoreplots: 1-2, 2-3, 1-3 and 1-4.

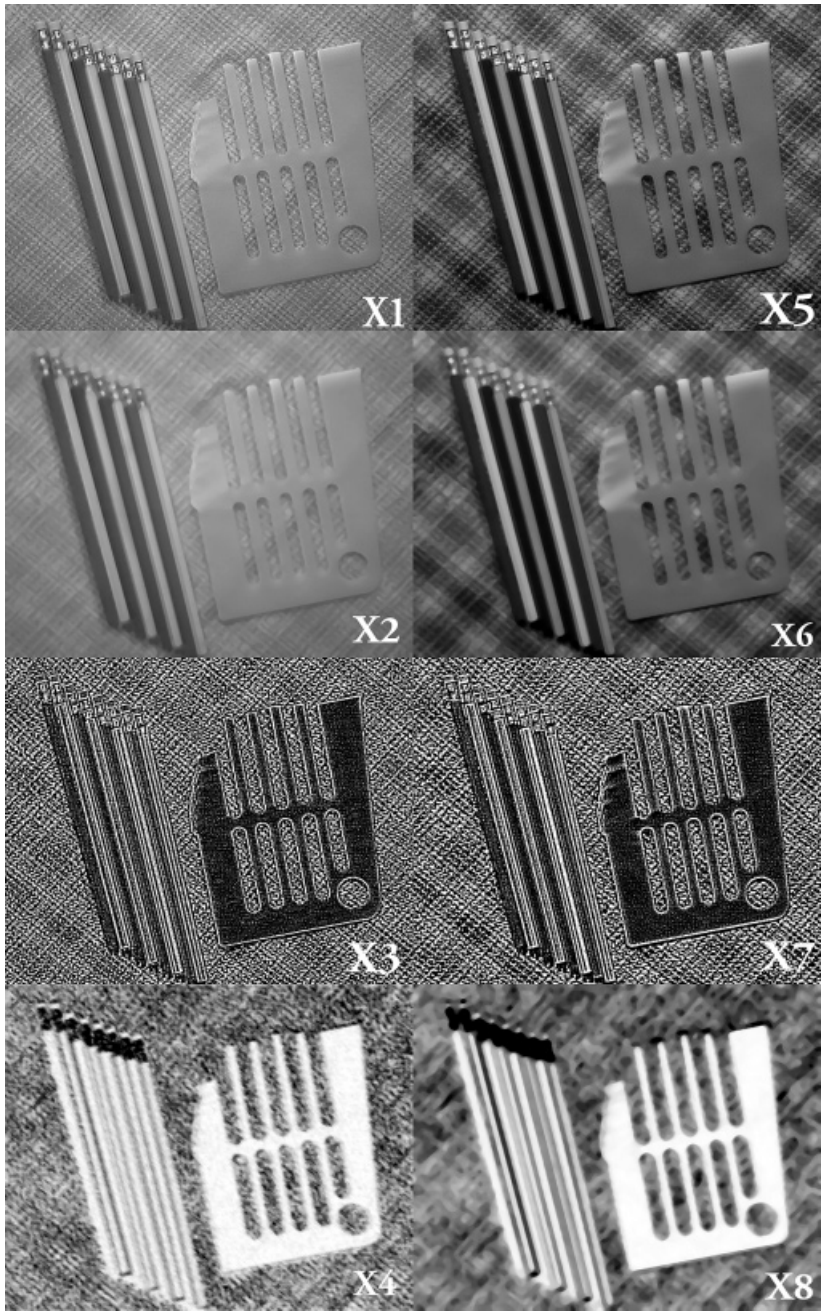


Figure 5. Two spectral channels (red, top left; green, top right) and three textural derivatives of each.

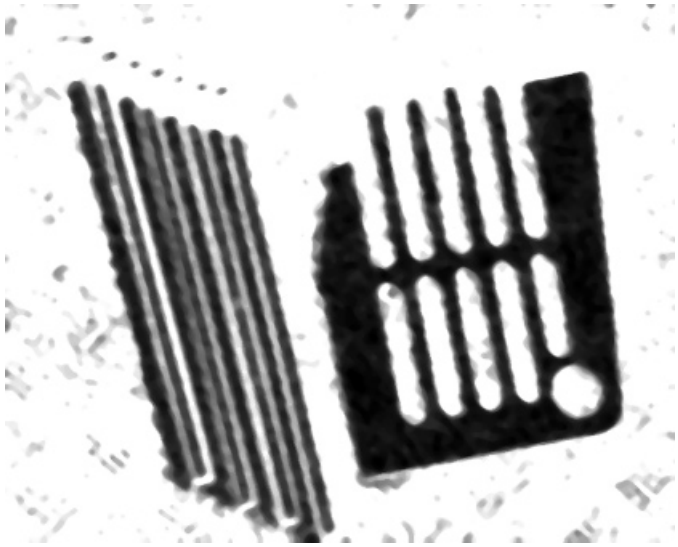


Figure 6. Reference **Y**-image expressing **TI** of the principal 'texture objects' ≠ spectral objects.

Table I shows the contents of **X** and **Y** in the three cases.

When comparing the three cases, the preprocessing must of course be identical. Thus, prior to the calculation in this example, all pertinent images were autoscaled.

Cases 1 and 2—MIA_0 and MIA_Y. In the first case, **X** contains eight variables and approximately 350 000 objects (pixels). Thus $\mathbf{X}'\mathbf{X}$ is an 8×8 matrix. In the second case, **TI** will be added to **X** from MIA_0 as an extra variable. Thus **X** will have nine variables and approximately 350 000 objects and $\mathbf{X}'\mathbf{X}$ will be a 9×9 matrix. The two models prove to be very similar, so loading plots are shown only for case 2. Score plots, though, are shown for both cases.

Figure 7 shows the accumulated explained variance for case 1. The number of PCs to use in the following discussion is not obvious from this plot, but using the standard four components that the software provides seems to be a fairly good alternative.

There is a very strong pairwise correlation between variables in these two cases. This can be seen from the loading plots (Figure 8).

One can see the following variable pairs in the loading plots: 1–5, 2–6, 3–7 and 4–8. An obvious interpretation would be that the texture filter operations on both red and green are closely similar. From these observations, one could for example argue that the number of variables could be reduced to four in the **X**-matrix, e.g. variables 1–4. If the computer is low on memory or speed, this can be considered to speed up the calculations. In the following, however, all the initial variables are used, since we have a quite different purpose than variable selection with the present decompositions.

Case 3—MIR. In the last case a regression model between the **X** used in case 1 (MIA_0) and **Y** from **TI** will be built using the kernel PLS algorithm. In this case the model will be actively forced in the direction of textural information, presumably somewhat suppressing pure spectral correlations, characterising the MIA_0 and MIA_Y cases respectively.

Table I. Contents of **X** and **Y** in MIA and MIR calculations

Channel	MIA_0	MIA_Y	MIR
Red	X1	X1	X1
Red median	X2	X2	X2
Red Laplace	X3	X3	X3
Red compound	X4	X4	X4
Green	X5	X5	X5
Green median	X6	X6	X6
Green Laplace	X7	X7	X7
Green compound	X8	X8	X8
Reference **TI** image (Figure 6)	—	X9	Y

In MIR there is still a correlation between the same variable pairs (see Figure 9), but not at all as strong as in MIA. The score image (see Figure 12) shows better textural details than in the previous cases. Thus putting the **TI** image in **Y** successfully forces the algorithm primarily to enhance texture in the decomposition, as it 'should' considering the exclusive texture index nature of the **Y**-image.

In general, of course, it is to be expected that MIA and MIR will decompose the same data set (the same multivariate image) differently, provided that the pertinent **Y**-reference information indeed does add new information. It is interesting to see how these expected differences manifest themselves in the loading and score plots (Plates 2–4) of the present example. Combining the first three score images into 'false colour composites' is always a useful way to compare alternative decompositions

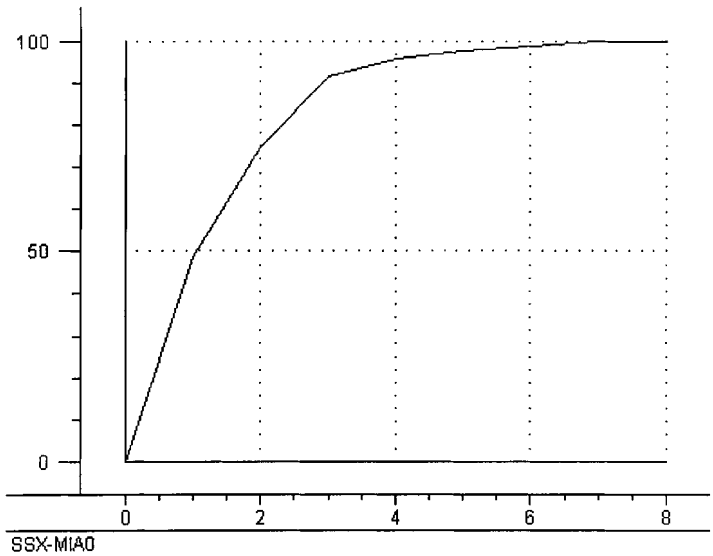


Figure 7. Explained variance for case 1, MIA without **Y**.

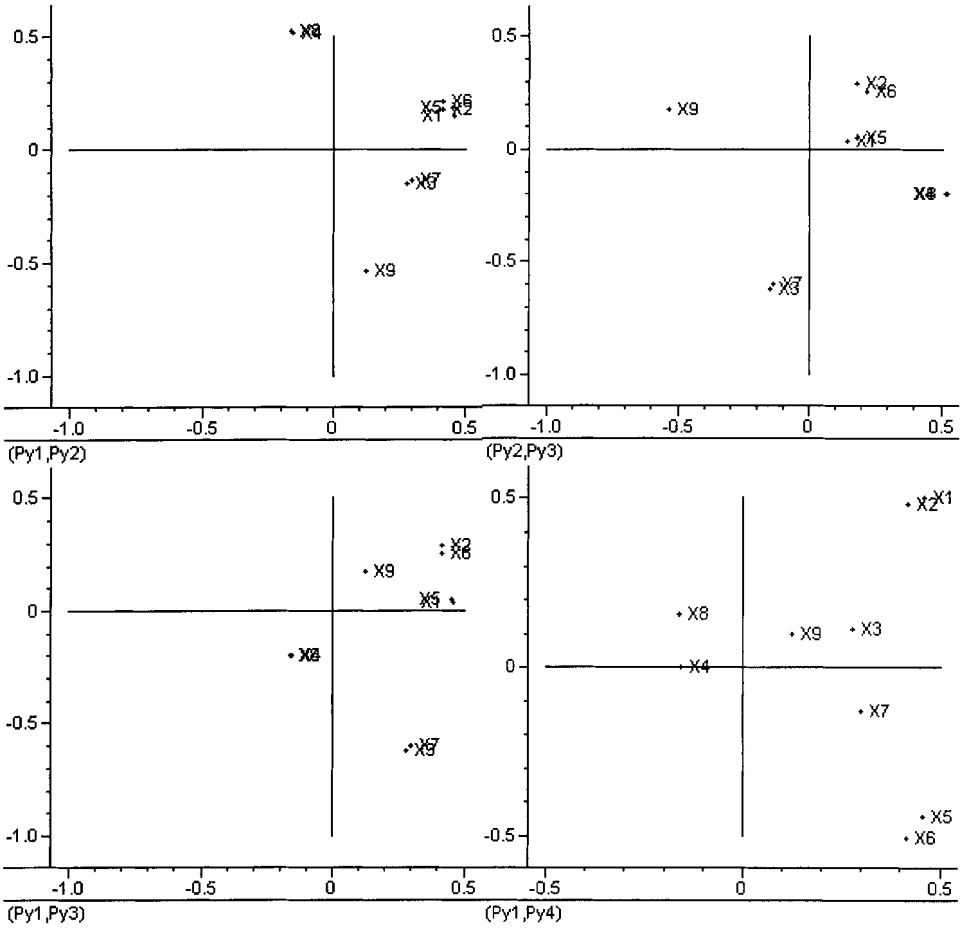


Figure 8. MIA_Y loadings 1–4.

(Figure 10–12). The most evident difference between the score images in these three cases, looking beyond differences in colour, is the gradual increase in detail. The MIR score image looks much 'sharper', more focused, than the MIA score images, primarily because of better texture and detail description.

Figure 13 shows the calibrated, explained variance for MIA_0 vs MIR [14,15]. It shows that, in this case, MIA performs better in the first two components than MIA. The third component is not very different in the two cases, while the fourth component is a little better in MIA than MIR. Figure 14 delineates y-variance modelled.

CONCLUSION

In this paper it has been shown that image PLS adds a new dimension to the complex field of analysing multispectral images. PLS was performed using the kernel algorithm, which is now implemented in our prototype MIA/MIR software system. The programming was done in a

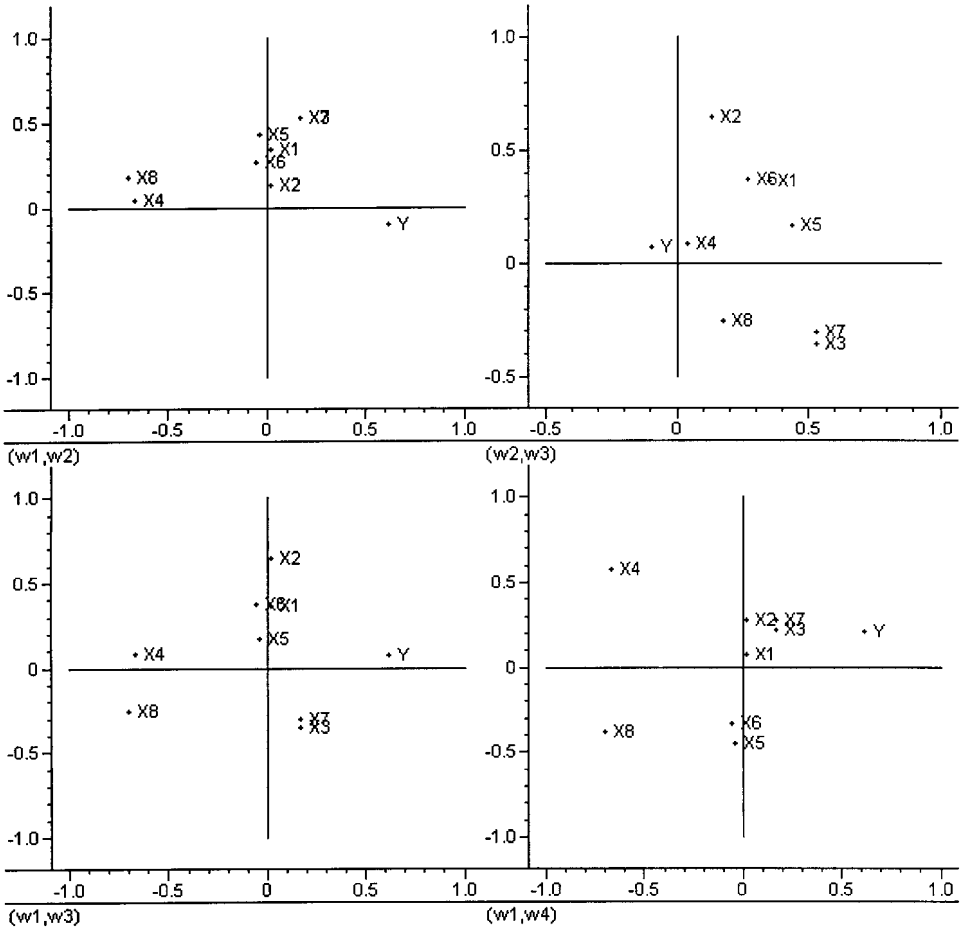


Figure 9. MIR loading weights 1–4.

combination of LabVIEW and MATLAB using the best properties of both programming environments. Using this approach, the calculations can be carried out on a standard desktop computer.

Applying this algorithm on a representative test image which shows many of the typical features found in technical imagery, we have shown that image PLS (MIR) decomposes the data differently than image PCA (MIA), in accordance with chemometric experience from ordinary two-way matrices. In the present example the **Y**-reference texture-related image used turned out to be able to force a rather significant 'tilting' compared with an 'ordinary MIA' of the primary structures in the original, spectral R/G image.

MIR requires a different validation approach than the conventional PLS approach. Much work remains, but the working prototype is now successfully implemented. We are currently also working on an extended series of representative applications.

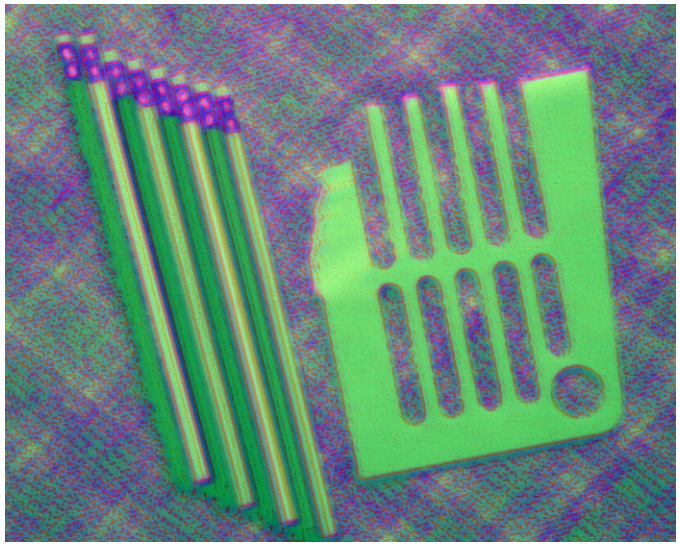


Figure 10. MIA_0 score images 1–2–3 (R–G–B).

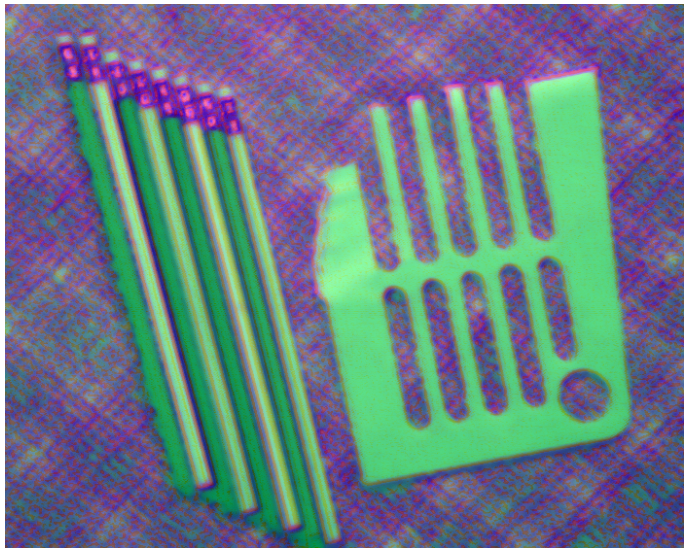


Figure 11. MIA_Y score images 1–2–3 (R–G–B).

Figure 12. MIR score images 1–2–3 (R–G–B).

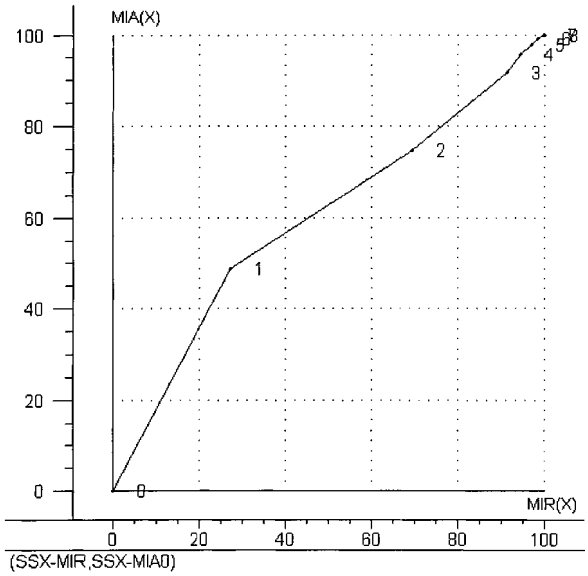


Figure 13. SSX MIA_0 vs SSX MIR.

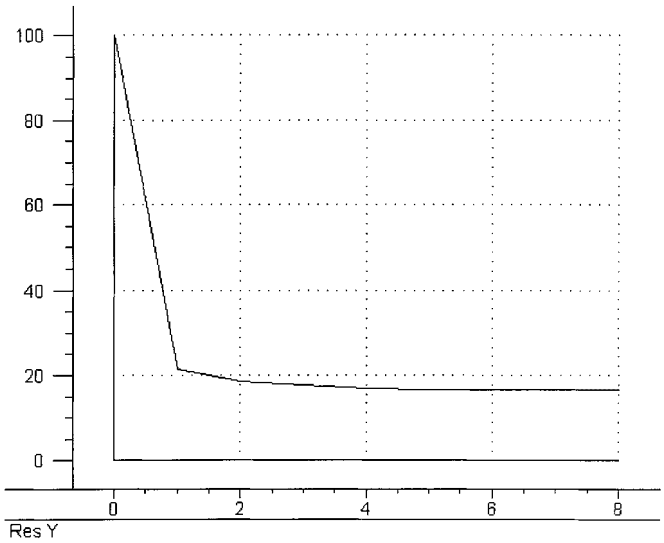


Figure 14. **Y**-residuals (SSY) from the MIR case.

REFERENCES

1. Esbensen K, Geladi P. Strategy of multivariate image analysis (MIA). *Chemometrics Intell. Lab. Syst.* 1989; **7**:67–86.
2. Lindgren F. Third generation PLS. *PhD Thesis*, Umeå University, 1994.
3. Alsberg BK, Remseth BG. Multivariate analysis of 2D surfaces—folding and interpretation. *J. Chemometrics* 1992; **6**:135–150.
4. Geladi P, Grahn H. *Multivariate Image Analysis*. Wiley: Chichester, 1996.
5. Esbensen K, Geladi P, Grahn H. Strategies for multivariate image regression (MIR). *Chemometrics Intell. Lab. Syst.* 1992; **14**:357–374.
6. Geladi P, Esbensen K. Regression on multivariate images. Principal component regression for modelling, prediction and visual diagnostic tools. *J. Chemometrics* 1991; **5**:97–111.
7. Grahn H, Szeverenyi N, Roggenbuck M, Geladi P. Tissue discrimination in magnetic resonance imaging. A predictive multivariate approach. *Chemometrics Intell. Lab. Syst.* 1989; **7**:87–93.
8. Grahn H, Sääf J. Multivariate image regression and analysis. Useful techniques for the evaluation of clinical magnetic resonance images. *Chemometrics Intell. Lab. Syst.* 1992; **14**:391–396.
9. Lindgren F, Geladi P, Wold S. Kernel-based regression: cross-validation and applications to spectral data. *J. Chemometrics* 1994; **7**:45–59.
10. Martens H, Næs T. *Multivariate Calibration*. Wiley: Chichester, 1989.
11. Rännar S, Lindgren F, Geladi P, Wold S. A PLS kernel algorithm for data sets with many variables and fewer objects. 1. Theory and algorithm. *J. Chemometrics* 1994; **8**:111–125.
12. Dejong S, Terbraak CJF. Comments on the PLS kernel algorithm. *J. Chemometrics* 1994; **8**:169–174.
13. Höskuldsson A. PLS regression methods. *J. Chemometrics* 1988; **2**:211–228.
14. Geladi P. The regression model comparison plot (REMOCOP). In *Frontiers in Analytical Spectroscopy*, Andrews D, Davies A. (eds). The Royal Society of Chemistry: Cambridge, 1995; 225–236.
15. Geladi P, Swerts J, Lindgren F. Multiwavelength microscopic image analysis of a piece of painted chinaware. *Chemometrics Intell. Lab. Syst.* 1994; **24**:145–167.

Paper III

Principles of MIR, Multivariate Image Regression - I: regression typology and representative application studies

by

Thorbjørn T. Lied & Kim H. Esbensen

	Corresponding Author	Senior Author
Name	Thorbjørn Tønnesen Lied	Kim H. Esbensen
Phone	+47 35 57 51 53	+47 35 57 51 50
e-mail	thorbjorn@lied.no	Kim.Esbensen@hit.no
Company	Telemark University College (HIT) Department of Technology (TF) Applied Chemometrics Research Group (ACRG)	
Address	Kjølnes Ring 56 N-3914 Porsgrunn Norway	
Fax	+47 53 57 52 50	

CONTENTS

Abstract 1

1.0 Introduction 3

1.1 Concepts 3

1.2 Software 5

2.0 Applications 6

2.1 IPLS-$Y_{discrim}$: Discrimination Prediction. 6

2.2 IPLS-Y_{grid}: Monitoring and estimating storage time for fruit (bananas) 12

3.0 Discussion and conclusions 18

References 20

Principles of MIR, Multivariate Image Regression - I: regression typology and representative application studies

ABSTRACT

We present an introduction to MIR: Multivariate Image Regression with a selection of illustrative application studies. Generalisation from 2-way multivariate calibration to the 3-way regimen leads to - at least - three alternative *image regression cases* depending on the nature of the available Y-data: IPLS-$Y_{discrim}$; IPLS-Y_{grid} ; IPLS-Y_{total} . A systematic *image regression typology* is briefly introduced.

We here present the core of the principles of *applied* MIR: Two major MIR application studies are worked through, a food mass product industrial inspection study (IPLS-$Y_{discrim}$) and a food product (fruit) storage stability image analytical monitoring (IPLS-Y_{grid}). These exemplifications are presented as *archetypes*, representing a much wider range of potential industrial/technological application areas.

The present paper represents one major element of our work towards establishing a complete, stand-alone facility for MIR (Multivariate Image Regression); the second paper in this series deals with the development, implementation and extensive exemplifications of a complementary cross-validation facility.

KEYWORDS: multivariate image regression; MIR; multivariate image analysis; MIA; multivariate image texture analysis; MIX; 2-D images; 3-D image arrays; *image regression cases*; applications

1.0 INTRODUCTION

The introduction of the Multivariate Image Analysis (MIA) concept in cheometrics was not longer ago than Esbensen & Geladi (1989) [1]. In the intervening 10+ years the development of MIA has been relatively slow, but would appear to begin to take to speed more recently - the entire field was summarised in the comprehensive textbook by Geladi & Grahn [2]. Much of the theoretical background for Multivariate Image Regression (MIR) necessitates a thorough understanding of the principles and methods in MIA, which we shall here assume known. It is especially important that the concept of the *multivariate image* is well understood.

Because regression calculations on the extensive amount of data in multivariate images easily can be a technical challenge (growing quadratically with the number of variables, or channels), some important recent evolutions have made it more feasible in the last few years. The continuously developing technical improvements in computer hardware efficiency is of course a major contributor, but especially the KERNEL PLS algorithm introduced by Lindgren et.al. in 1993 [3] has dramatically speeded up the present type of calculations, as was outlined in detail in [5].

Despite of this, few studies has yet shown the true potential of MIR as a tool for predicting quantitative features in multivariate image data. Hopefully, this condition will be improved by the current paper.

1.1 Concepts

Several concepts are used in this paper, some of which may be relatively unfamiliar. A brief introduction of these is given to help the reading of the article.

1.1.1 Multivariate Image (MI): The MI is a digital image of *one* scene, consisting of *many* variables (channels), e.g. colour bands, channels. At the outset the simplest situation is the one in which each image *pixel* is treated as an *object*, which requires rigid consistency in scene lay-out for all variables. An object in a given scene position in one variable must be found in the same scene position in all other channels; for regression cases also in the Y-image. MI's are usually presented as a 3-D matrix, but because the two object-ways can be treated as one way, the MI may also be reorganized into a 2-D-matrix prior to modelling, and 2-way methods can be applied [2] by way of the so-called *unfolding* operator.

1.1.2 MIR: Multivariate Image Regression [2, 4] builds regression models between the multivariate X-image and the (uni-/multi-)variate Y-image. MIR is here performed

using KERNEL-PLS [3, 5] on reorganized Multivariate Images, i.e. each variable is reorganised into a (very long) object vector. In this basic unfolded form, MIR uses only the variable-signatures, i.e. the spectral information in the analysis and only indirectly makes use of the spatial information analogous to MIA [1, 2]. But even though MIR technically uses 2-way analytical methods, there is an enormous visualisation potential in image data which is also used fully in MIR. Displaying results not only in score space, but also in the so-called *backfolded* image space, enhances the insight in the data structure and the models developed. Applying colour coding to score plots, MIA, or by combining three score images in one composite ("R/G/B") colour image, it is often possible to capture comprehensive model presentations of great interpretation value etc..

1.1.3 MIX: Multivariate Image teXture analysis. MIX is an extended MIA-MIR approach which includes spatial, especially textural, information in the analysis. In cases where spatial information is important, this can be included in the MIR-model by e.g. adding derived textural variables calculated from the original variables [6-8]. Sometimes enhancing details using e.g. edge-detectors is favourable, in other cases the opposite (smoothing details) might be required etc. Also, *combining* textural filters might often give very useful results. MIX has the potential for explosive data growth, thus powerful means for variable selection are required. We do not treat the MIX aspects in any depth in the present work however.

1.1.4 Regression typology

Perhaps surprisingly, going from the 2-way realm in which the concepts of multivariate calibration is well-known - and need no further presentation here - there is a corresponding three-fold multitude of analogous but in principle different image *regression modes*, Figure 1.

1.1.4.1 IPLS-$Y_{discrim}$: The Yes/No classificator/discriminator. In every position in the Y-image, a pixel is either 1 (one) if it is part of a current class, otherwise 0 (zero). The approach is suitable for classifying one class among (many) others. Used as a pre-processor, this method can easily be taught how to pick out desired classes. This case is also easily extended to cover several classes, by using several one-class Y-discrim masks, Figure 1.

1.1.4.2 IPLS-Y_{grid}: If every Y-condition is not available in one image, several images can be juxtaposed in a compound, so-called *gridded* image. This way the total experimental design can be represented in one image, i.e. one model. Extensive

illustration of IPLS-Y_{grid} is given in this work. In some cases, especially when predicting an overall value for each sub-image in the grid, the corresponding Y-image will have a constant value within each sub-image. When this appears, some kind of smoothing of each sub-image in X will usually be useful, i.e. reducing non-classification related variations in X.

1.1.4.3 IPLS-Y_{total}: When the entire experimental design is covered in one frame, merging images together, as in the IPLS-Y_{grid} is not required. In these cases, each pixel in X also has a separate, unique value in the Y-image. Typical examples come from e.g. remote sensing. Because most of the still limited MIR-literature explicitly discusses this kind of data, and because it is merely a special, extreme case of the Y_{grid}, it will not be treated further in this paper

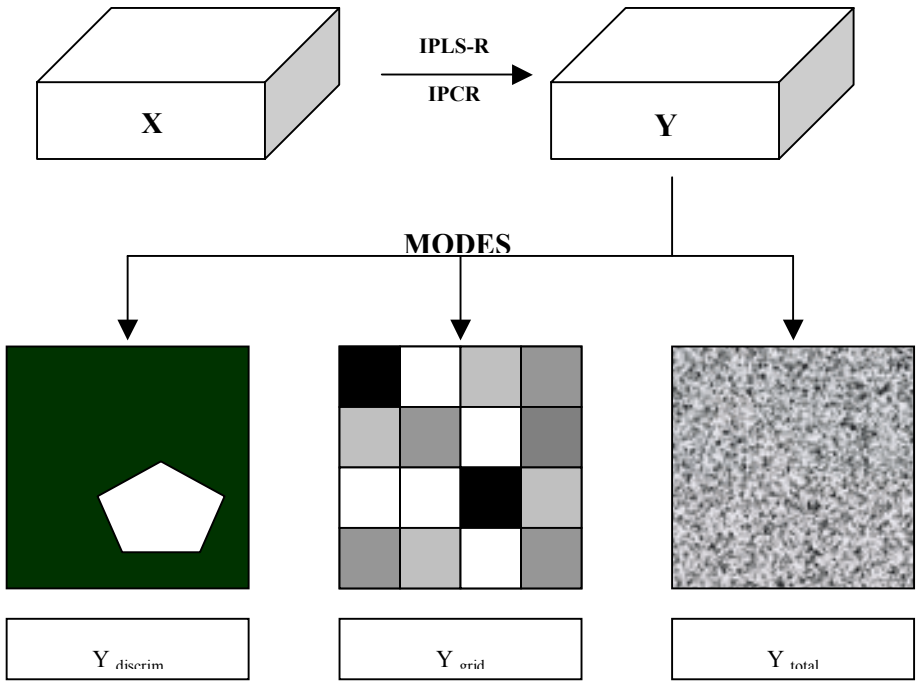


Figure 1. The three different MIR modes, $Y_{discrim}$, Y_{grid} and Y_{total}.

1.2 Software

All calculations in this paper are performed using a self-developed program, described in Lied, Geladi & Esbensen [5]. The software is available for Microsoft Windows ® (9X NT 4+5) and is written in National Instruments' LabVIEW v. 5.1. Both MIA and

MIR is implemented; for MIR regression calculations, KERNEL-PLS [3] is used exclusively. [1]

2.0 APPLICATIONS

Below the terminology IPLS (Image PLS) is used throughout, but it is evident that PCR *may* also be used alternatively should one so desire, albeit with the well-known distinctions regarding PCR vs. PLS [9-10] etc. Here PLS is employed exclusively because of its well-known chemometric advantages [9-14]. In both major examples below, data are mean centred and scaled to uniform standard deviation. All variables thus have equal variance weights, making the Kernel-PLS decompositions pertain to *correlations*.

2.1 IPLS-$Y_{discrim}$: Discrimination Prediction.

Motivation: A *pilot study* of image analytical *industrial inspection* of a mass production food article, Swedish crispbread ("knäckebrød") is presented. This item, by nature of its mass consumption status, is produced in very large quantities in industrial bakeries in many countries. Output from the industrial ovens are necessarily way outside *complete* human inspection capabilities, for which reason an automated, industrial image analytical system would be of considerable interest. This in turn could form the basis for a truly 100% inspection system.

In our restricted pilot study involving some 10 pieces of crispbread, parallel representative oven outputs are available, 5 with an "accepted" status and 5 with three types of representative faults, typically encountered in the industrial production situation. Figure 2 shows these three faults together with examples of the directly acceptable product ("normal"). Technical details regarding this image is found in table 1.

---

[1] When developing this prototype, serious efforts were made to enhance the flexibility and user interaction facilities. For large datasets, 10 M pixels or above, calculations starts to become slow however. Development of a professional system is now under way. Contact the corresponding author for instructions on how to download the freeware *prototype*.

Table 1. Technical Details of the crispbread case

Image Capture	**Camera**	**Lens**	**Focal Length**
	JVC 3CCD KYF-50	**Micro Nikkor AF**	**105mm**
Measures	**With (pixels)**	**Height (pixels)**	**# Variables**
Total Image	1000	666	4
Sub Images	200	333	4
Spectral Variables	**Colour**	**Wavelength**	**Bandwidth**
1	RED	N/A	N/A
2	GREEN	N/A	N/A
3	BLUE	N/A	N/A
Textural Variables	**Filter**	**Window Size & Passes**	**Applied to**
1	Variance		

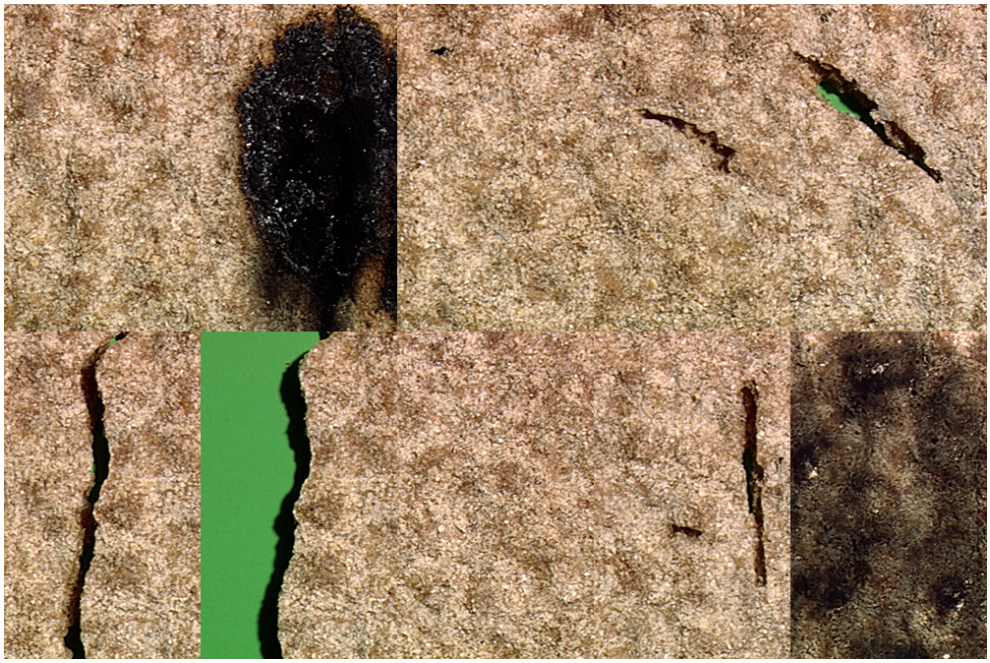


Figure 2. "Normal" and "flawed" Scandinavian crispbread ("knekkebrød"). Three representative types of flaws are displayed; broken, perforated and burnt cases.

Figure 3 shows standard MIA score plot set-ups for the crispbread case (figure 2) [1, 2].

In the interest of the wider application horizon for this specific IPLS-$Y_{discrim}$ regression example, a more *general* IPLS-$Y_{discrim}$ regression case will be set up. This is done by using a *Y-image mask* of the areas of interest in the image which are recognised as "rejects", i.e. areas which are underlain by those parts of the Y-image which depicts flaws of the various type. Figure 4 shows this "flaw"-mask.

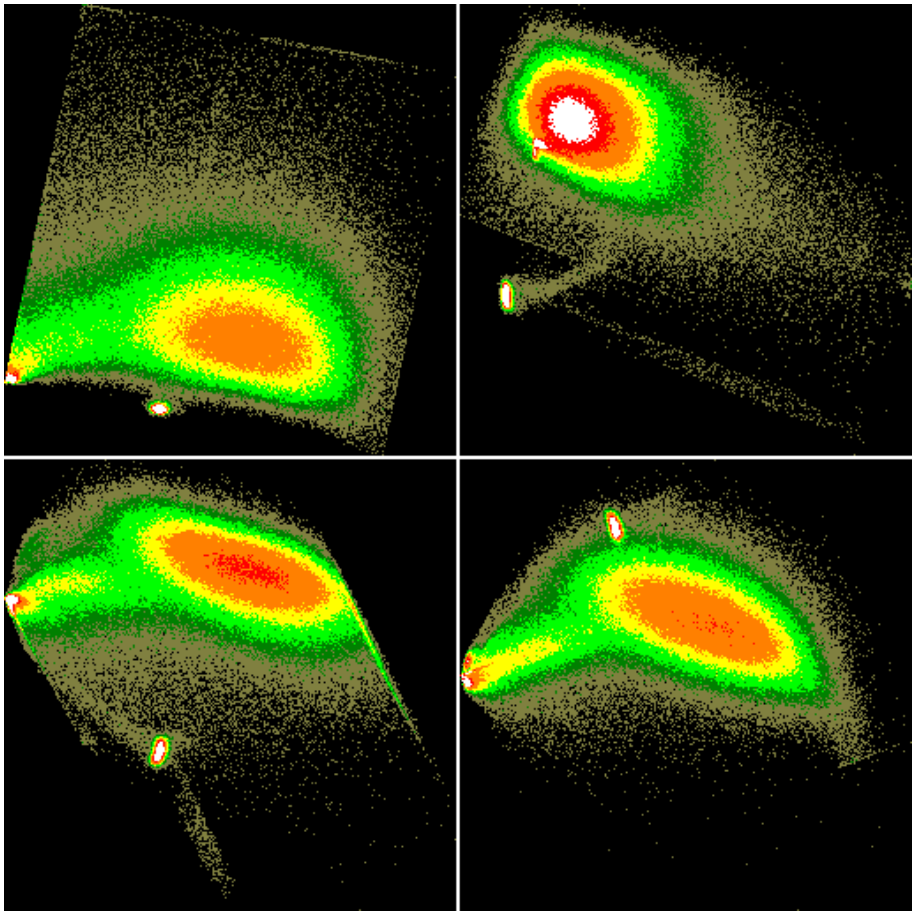


Figure 3. MIA standard score-plot for the crispbread case in figure 2.

Figure 4. Y-image mask of (0/1) discrimination areas. Note that by using a relevant background discrimination feature, it is possible to zoom in only on the true flaws present in the gridded calibration imagery, which have been designated white here.

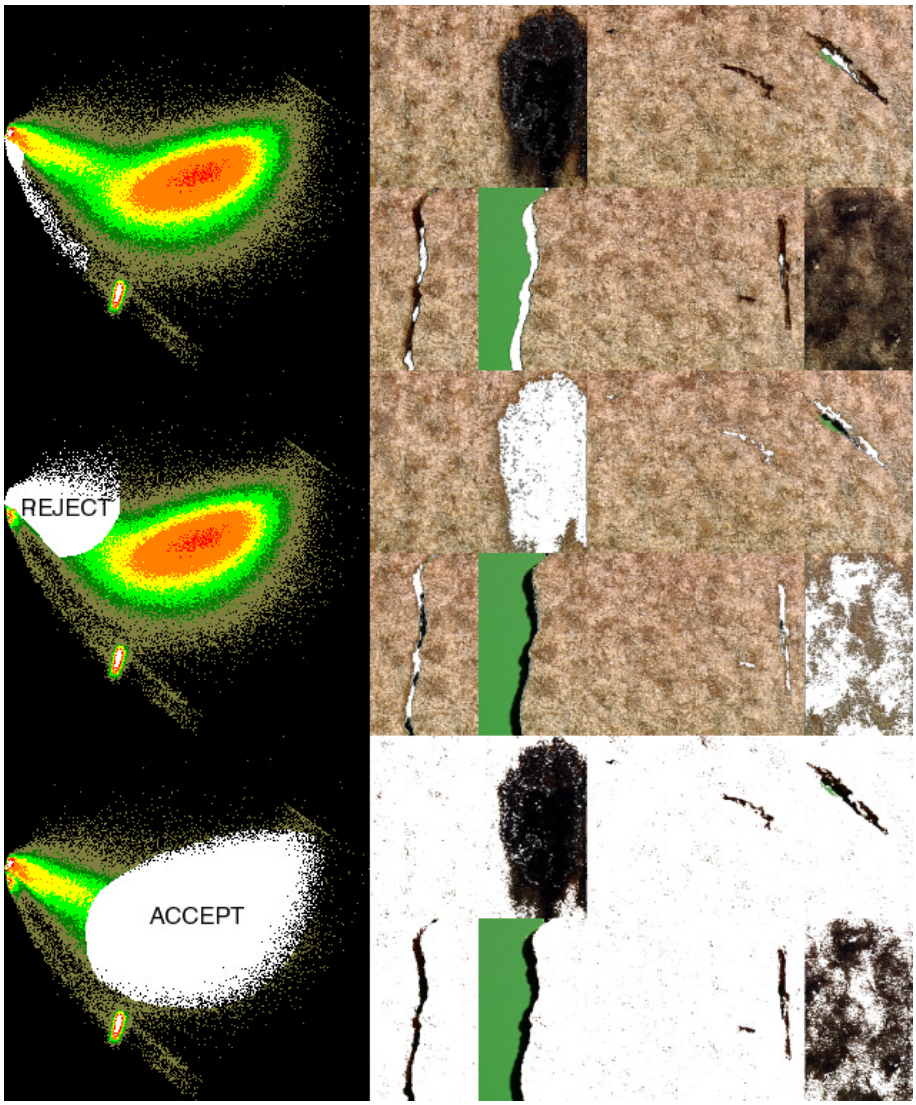


Figure 5. IPLS-$Y_{discrim}$ t1-t4 score plots showing all three resolved classes in the crispbread case: broken/perforated (top panel); burnt (middle) and "accepted" (lower). Note complete discrimination. Also comp. with similar t1-t2 scoreplot from the MIA-solution

Figure 5 has been designed to bring forth the full potential of the IPLS-$Y_{discrim}$-case, showing (in standard MIA-style) the corresponding t1-t4 score plots *versus* the original (raw) image domain layouts of three fault classes present in the fully background-discriminated crispbread case: shadow (top); burnt (middle) and "accepted" (bottom). It is gratifying to observe *complete discrimination* between all relevant classes, i.e. all three types of rejects/accepted and the background as well. This

successful discrimination points directly to the desired use of image-based *prediction* of all these types crispbread. This pilot study, while extremely simple, allows full conceptual delineation of a complete automated image analysis system, by way of the relevant PLS-prediction facility, Figure 7.

Figure 6. Calibrated X-variance (top) and Y-variance (middle) and validated PRESS (low) for the crispbread case in figure 2.

Incidentally, observe that in this particular case, there would appear to be very little "tilting" of the IPLS-solutions relative to the corresponding simpler PCA-solutions (MIA), contrary to many other two-way experiences [4-5, 15-16]. In the present case this reflects a rather direct correspondence of the X-block data structure with the Y-structure(s), i.e. the information gathered in the image analytical X-decription "happens" to be directly correlated to the guiding Y-discrimination dummy variable; see also below however.

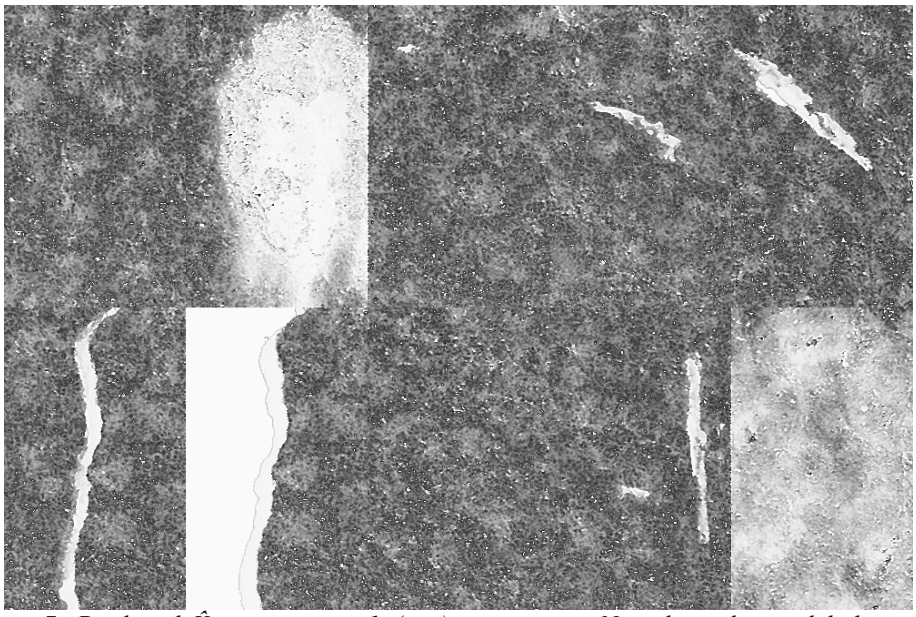


Figure 7. Predicted Ŷ-image using 1 (one) component. Note how the model distinguishes clearly between faults/non faults.

While figures 4-6 gives the statistical facts in this case, figure 7 shows the actual predicted image in scene space using one component. This figure demonstrates that the model is excellent for predicting all the relevant types of faults. This result is, needless to say, of a much larger generalisation potential than the specific crispbread example chosen. The illustration in fact has merit as an *archetype* for IPLS-$Y_{discrim}$ multivariate image regression.

MIX-aside: In this example a variance filter (table 1) has in fact also been applied to extract local textural variations in the X-image. While the background is flat, the crispbread has a very distinct, regular texture. The variance filter, which returns the local variance in a small window in every position in the image, will thus greatly help to *distinguish* crispbread from background as well as textureless burnt parts, assisting

the spectral information in the classification. The effect of this filter addition is visualized in figure 8 which shows the *loading weights*(w) for this example. From this figure, it can also be seen that X2 and X3 (Green and Blue) contains mostly the same information, indicating that one of them could be left out in later calibrations etc.

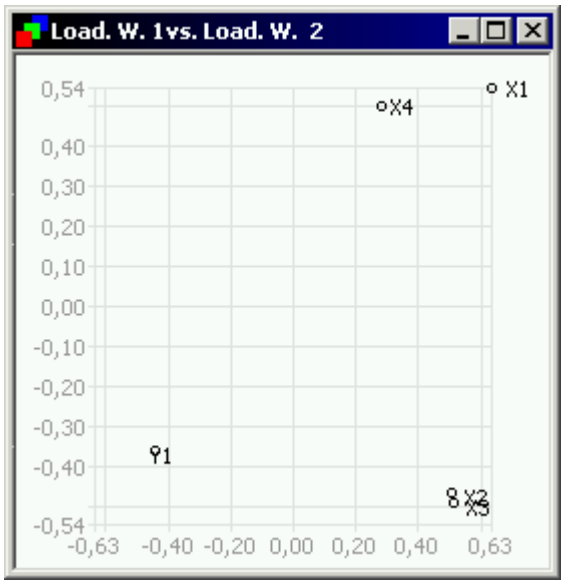


Figure 8. Loading weights 1 vs. loading weights 2 for the crispbread case. Note X-variables 2 (green) and 3 (blue) which seem redundant.

Note how judicious use of a *relevant background* (colour, texture) is essential to bring about the successful discriminations in this case. As in all image analysis applications *illumination* and *colouring* (not treated here) is often of equal importance compared to the data analysis proper etc.

2.2 IPLS-Y_{grid}: Monitoring and estimating storage time for fruit (bananas).

Motivation: The objective of this application example is to monitor storage stability by a series of (multi-temporal) images of the same fruit(s), with great efforts to keep all storage and imaging parameters *constant*, the only variable being time elapsed since storage start. Successful monitoring will allow for quantitative storage deterioration prediction directly from the captured multivariate X-images [5].

In this context, the calibration-parameter "storage time" shall be represented by juxtaposed part-images, making up a complete, so-called *gridded*, multivariate image, hence the suggested name for this second image regression mode: IPLS-Y_{grid}. This

example also serves as an *archetype* not only of multi-temporal studies but also of analogous objectives, conf. below.

Figure 9 shows the compound, gridded multivariate X-image of a deteriorating banana, for which the storage times involved are (from upper left to lower right): 1/2/3/6; 7/8/9/10; 13/14/15/20 days respectively. This gridded layout is necessary in order that all storage times can be *analysed together* by MIA or MIR. It is emphasized that it is the *objective* of the image analysis (in this case: storage stability monitoring) which *dictates* that the individual grids represent a succession of different storage times. For other image regression cases, these individual grid cells will often represent different, typical "object-like" categorical entities to be similarly *compared*, e.g. a series of different meats to be characterised, as was the case with Wold et al. [15] (in fact also predicting a storage-related parameter, "harshness"), Geladi and Grahn [2] a.o.

Table 2. Technical Details of banana example

Image Capture	**Camera**	**Lens**	**Focal Length**
	SILVACAM	**Fujinon**	**120 mm**
Measures	**With (pixels)**	**Height (pixels)**	**# Variables**
Total Image	800	600	3
Sub Images	200	200	3
Spectral Variables	**Colour**	**Wavelength**	**Bandwidth**
1	NIR	760 - 900 nm	
2	Red	580 – 680 nm	
3	Green	490 - 580 nm	

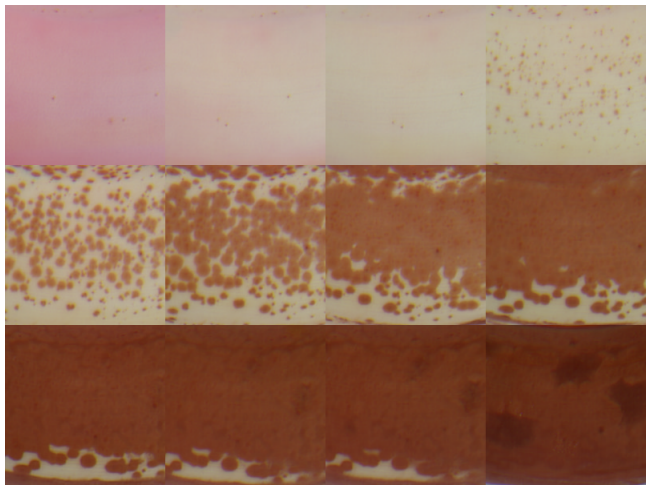


Figure 9. Storage time aging of representative fruit (banana) in the interval 1-20 days.

In figure 10, which shows the particular *Y-image*, the *array* of grid cells forms the basis for an IPLS1. Observe how the deterioration process interval of 1-20 days has been *mapped* into an image analytical appropriate grey-level interval, spanning 0-255. Again it is the Y-image mask which makes the regression problem immediately appreciable. In one sense, as soon as the multivariate X-image has been defined, it is the Y-image which sets up the entire MIR.

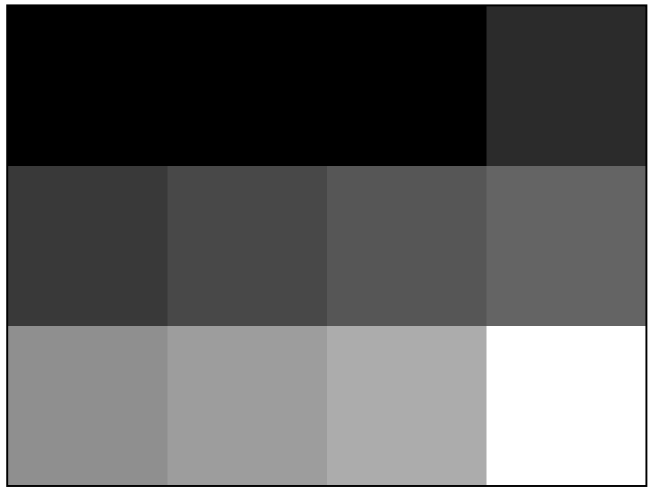


Figure 10. IPLS-Y_{grid} Y-image, delineating the problem-dependent Y-levels for the banana deteriorating process. Figure 10 (Y-image) corresponds to figure 9 (X-block).

Figure 11 performs an identical role as figure 5, encompassing the essentials of the IPLS-analysis. In the t1-t2 score plot (upper left panel) one may appreciate, in full detail, the *trace* of the fruit deterioration process. [2] We have illustrated two representative *process stages* along this trace, an intermediate stage and the penultimate sad, almost totally rotten end of the banana development (upper right and lower left panel respectively). With reference to MIA [1, 2, 17] the scene-space back-

[2] We have elsewhere worked out a complete image analysis strategy, which - while originally presented as related to MIA - also applies to the *analogous* t-t- score plots derived by an IPLS-solution [17]. As but an example we there followed another biological process, albeit of considerable greater complexity, i.e. a forest clearing *regrowth* process, using geomorphological analogies in order to characterise MIA score plots. From this review [17] a range of interpretation guidelines for t-t score plots were developed, all of which may also be applied to the present PLS-solutions. Observe that a slightly different modus operandi applies to the t-u plots [18]. These subtle differences will be addressed in several sequel papers on a comprehensive MIR strategy, which are in the works.

projections of these two classes are self-explanatory in figure 11, especially when compared with figure 9. Fig. 11 shows how it is possible to delineate the *entire* deteriorating process in the X-space because the *entire* storage time *calibration span* has been compounded in the one (X,Y)-image.

For IPLS solutions the t-u score plot allows valuable, indeed critical insight into the *effective* regression relationships between the X- and the Y-space [18]. E.g. if already the t1-u1 relationships is (close to) linear, this is a certain reflection that a strong prediction model will be achieved; likewise, smaller non-linearities in the t1-u1 score plots are usually "ironed out" by inclusion of one or a few, *additional* PLS-components t2-u2, t3-u3 etc.

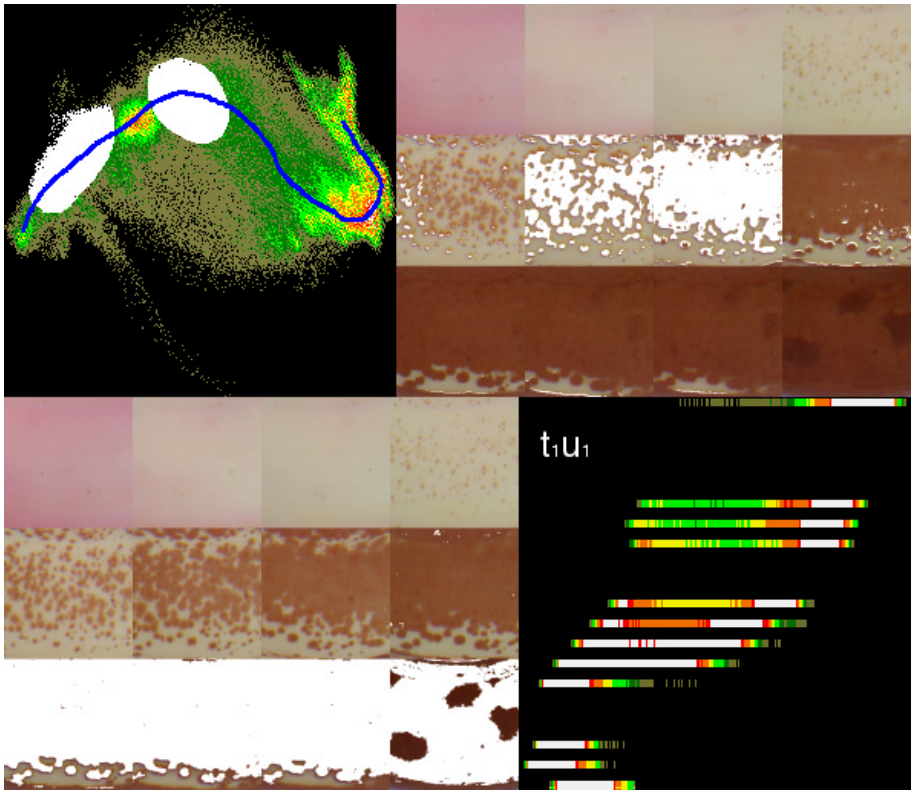


Figure 11. IPLS of the fruit aging process, in the IPLS-Y_{grid} regression case. Upper left: t1-t2 score plot, with two sets of corresponding scene-space (upper right and lower left panels). Lower right: corresponding t1-u1 score plot. Note discrtesation along the u1-axis, corresponding to the Y-levels presented in figure 9.

For the present first presentation of the most used features in multivariate image regression, these few aspects of the general use of the t-u- plot will be enough to allow

appreciation of the way the IPLS-Y_{grid}-*modelling* works. Figures 11-13 represents salient central aspects of our work leading up to a complete *MIR strategy* [5,19].

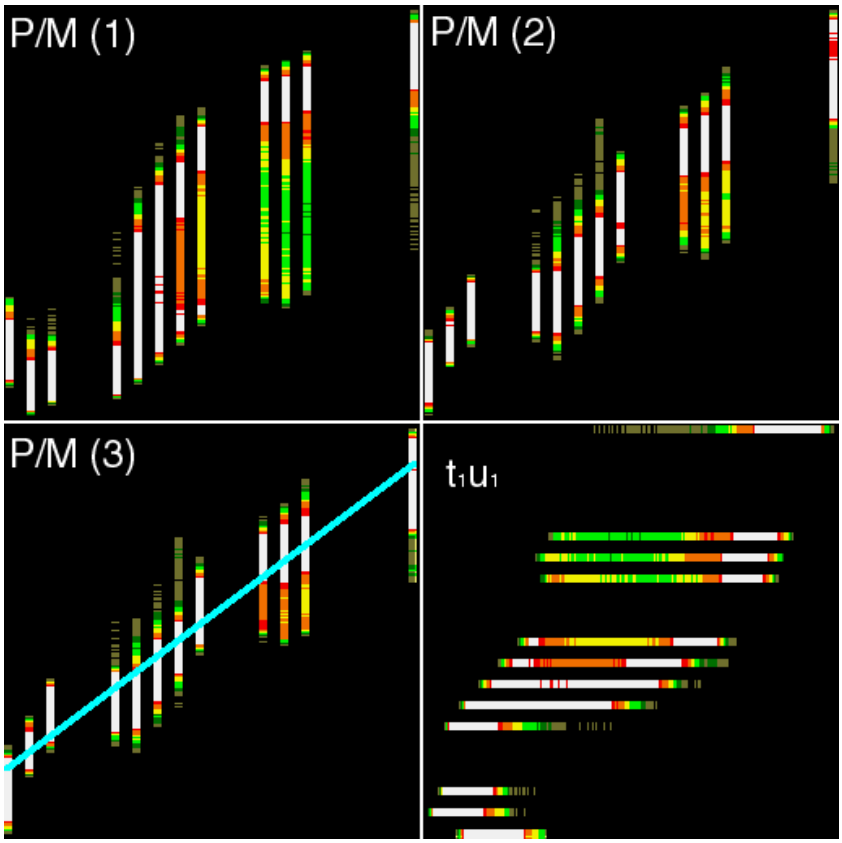


Figure 12. Complete predicted vs. measured (P/M) layout of the banana aging process IPLS-Ygrid-analysis. The standard P/M assessment plot is shown for 1,2 and 3 IPLS-components, while only the t1-u1 scoreplot is shown (lower right). Significant improvement of prediction precision using 3 components

For an assessment of the *modelling strength* of the IPLS-analysis, figures 11-13 will also suffice. From these X-Y relationships it is evident that a satisfactory model can be achieved using three IPLS-components. Observe e.g. how the P/M (*predicted* vs. *measured*) relationships improves quite considerably when adding the second - and third IPLS-components. From the t1-u1 relationships *alone* it was however already clear that this would *per force* result. We are also able to follow how one would go about identifying *outliers* etc. by using the appropriate t-u score plots, following [18]. In the specific present plots in Figure 11 we did actually not have reason to perform any outlier deletion, since none were found.

Figure 13. Calibrated X-variance (top) and Y-variance (middle) and validated PRESS (bottom) for banana aging case.

3.0 DISCUSSION AND CONCLUSIONS

The examples above represents our first presentation of a *simulation* of an automated image-analysis monitoring system in the guise of IPLS: MIR (and MIX), in which we focused on the general aspects of the IPLS-$Y_{discrim}$ and the $IPLS_{grid}$ *cases*.

The specific choices of illustrative food article systems is not in any way an absolute indication of the general applicability of this approach. What have been shown feasible for perishable fruit articles, and for on-line food product characterisation, is of course equally applicable to, say, cereals, bread, meat, fish - indeed the food and feed areas at large are potentially opened up for a similar approach, e.g. the berries -, beverages, - dairy sectors etc.

Continuing outside the human, and animal, food and feed areas, an *analogous* automated image monitoring approach can of course equally well be envisaged for quite different application areas, at first primarily within the general technological and industrial sectors - but perhaps even further removed. One common denominator could be any *multi-temporal* aspect, which would lend itself to an appropriately modified image recording and - analysis approach, *similar* to the one *illustrated* here, e.g. degradation studies: paints, coatings, corrosion inhabitation - industrial inspection in general i.a. Within the field of *remote sensing* there is also a plethora of similar multi-temporal objectives.

The on-line image monitoring example, while relatively simple in the crispbread case, also has many, much broader applications potentials within much of the *industrial inspection* realm, in which there is often a distinct need for automated image analytical monitoring.

For the present feasibility studies we are satisfied with the above results for both the *IPLS-*$Y_{discrim}$ and the *IPLS-*Y_{grid} approaches. We have shown that the multivariate image regression approach (MIR) is now fully established. It bears in mind though, that there is always a series of critically important *specific* associated image-analytical problems, e.g. problem-specific illumination, shadows, reflections, non-constant object sizes i.a. - Much interesting work remains.

The present first foray into the possibilities of multivariate image regression has focussed on the ways-and-means of *modelling* (using bilinear IPLS) and *prediction*. What remains is the equally important aspect of multivariate calibration, *validation* (in the form of image-regression validation), which forms the subject-matter of the second

paper in this series, in which we also will make use of the third IPLS-regression case only identified here: Y_{total} [19].

References

1. K. Esbensen and P. Geladi; Chemometrics Intell. Lab. Syst. **7** (1989) pp. 67-86.
2. P. Geladi, and H. Grahn: Multivariate Image Analysis. John Wiley & Sons, Chichester, UK, 1996, p. 316.
3. F. Lindgren, P. Geladi and S. Wold; J. Chemometrics, **7** (1993) 45-59
4. K. Esbensen, P. Geladi and H. Grahn; Chemometrics Intell. Lab. Syst**. 14** (1992) 67-86.
5. T.T. Lied, P. Geladi and K. Esbensen; J. Chemometrics **14** (2000) 585-598
6. N. Lamei, K.D. Hutchison, M.M. Crawford and N. Khazenie. Optical engineering **33** (1994): 1303-1313.
7. T. Yamazaki and D. Gingras. IEEE transactions on image processing **4** (1995) 1333-1339
8. J. R. Carr. Computers & Geosciences. **22** (1996), 849-865.
9. H. Martens and T. Næs: Multivariate Calibration. John Wiley & Sons, Chichester, UK, 1989, p. 419.
10. K. Esbensen, S. Wold and P. Geladi; J. Chemometrics, **3** (1988) pp. 33-48.
11. A.K. Smilde, J.A. Westerhuis and R. Boque; J. Chemometrics, **14** (2000), pp. 301-331.
12. R. Bro, Doctoral thesis. (1998): Multiway analysis in the food industry.
13. C. Anderson, and R. Bro (editors); J. Chemometrics **14** (2000): pp. 103-334.
14. A.K. Smilde, and H.A.L Kiers; J. Chemometrics **13** (1999) pp.31-48.
15. P. Geladi and K. Esbensen; J. Chemometrics **5** (1991) pp. 97-111.
16. J. P. Wold and K. Kvaal ; Appl. Spectrosc. **54**. (2000)
17. K. Esbensen, T.T. Lied, K. Lowell and G. Edwards. (Submitted for publication) Principles of Multivariate Image Analysis (MIA) in remote sensing, technology and industry.
18. K. Esbensen: Multivariate Analysis in Practice, 4th edition. (2000). CAMO ASA, Oslo, Norway.
19. T.T. Lied and K. Esbensen (Submitted for publication) Principles of MIR, Multivariate Image Regression -II: Cross validation- What you see is what you get.

Paper IV

Principles of MIR,
Multivariate Image Regression - II:

Cross validation - what you see *is what you* get

Thorbjørn Tønnesen Lied & Kim H. Esbensen

thorbjorn.t.lied@hit.no - kim.esbensen@hit.no
+47 35 57 51 53 - +47 35 57 51 50
fax: +47 35 57 52 50

Telemark University College
Dept. of technology
Kjølnes ring 56
N-3918 Porsgrunn
Norway

CONTENTS

Abstract 1

Introduction 1

Nomenclature 4

Case studies 4

Case 1: Full Y-image 7

Case 2: Problems 10

Case 3: Y-grid 11

Case 4: Cutting to the bone 17

Discussion and conclusions 19

References 20

ABSTRACT

This paper deals with generic problems regarding segmentation for cross validation in multivariate image regression. Multivariate images are characterized by a very large numbers of pixels which usually are highly redundant. When several thousand (ten thousand) pixels or more represent the same object, special considerations are required for proper cross validation segmentation.

A new approach for *guided segmentation* is introduced, in which the validation segments are specifically delineated by the informed user in score space. The practise of "blind", automated segmentation, which is dominating 2-way cross validation, is found to be useless in the 3-way MIA regimen. Problems concerning which order of components to use for the segmentation delineation are illustrated and the necessary precautions needed to ameliorate this approach are discussed. A general solution to the problem, called *higher-order components guided random sampling* , is described in detail, which may even also shed new light on current chemometric cross-validation practises in the conventional 2-way realm.

This new cross-validation approach is illustrated with multivariate image data sets which are known from the pertinent literature for easy comparison.

INTRODUCTION

This paper is the second in a series regarding Multivariate Image Regression, MIR, which has been developed to create regression models between multivariate images [1]. For a general introduction to this field, please see part 1 [2], in which the complete phenomenology of the three principal cases of multivariate image regression was detailed.

A multivariate image is a 3-D OOV matrix [3], i.e. two ways are objects (pixels in rows and columns), while the variable-way is comprised by different channels, e.g. colours. There are quite distinct differences between this 3-way domain and the complementary OVV domain, well-known from the three-way decomposition. These two domains do not in general make use of the same data modelling methods [2]. Here we treat OOV (MIA, MIR) exclusively.

In any multivariate model that will be used for prediction, it is important to know the predicting powers of the model. This is usually done by estimating the prediction errors as a measure between known and predicted values. A popular prediction measure is RMSEP (Root Mean Square Error of Prediction) which is defined as

Equation 1

$$RMSEP = \sqrt{\frac{\sum_{i=1}^{n} (\hat{y}_i - y_{y,ref})^2}{n}}$$

where $\hat{y}_i$ refers to the predicted value, and $y_{i,ref}$ is the known value [4].

The procedure of testing prediction performance is known as *validation*. To perform this optimally, at least two sets of data are required, one for calibration and one for validation. When a model has been established, using the calibration set, the validation set is subsequently used for predicting the ŷ-values of the validation set for comparison, e.g. according to equation 1.

At least two variations for this type of validation exist, one is known as "*test set validation"*, the other as *"cross validation"*. In test set validation, a completely new, independently sampled dataset is acquired, in addition to the calibration set. This demands that an identical sampling procedure is used for both data sets.

If this is not feasible, a different, less optimal, approach will have to be resorted to. *Cross validation* extracts a pseudo-validation set from the calibration set before building the model on the remaining complement of data. The extracted data is now used for validation. This approach may take several different forms, but all are closely related, in that they per force must correspond to a number of so-called segments in the list: 2,3,4,5....N, where N stands for the total number of objects in the original calibration set. After prediction errors have been estimated for the one left-out segment, it is replaced back into the modelling base and a new model is created in which a different segment is being kept out of the modelling etc. This is continued until every segment, and object, has been used for validated, hence the term *cross* validation [5].

To get realistic validation estimates, it is important that the calibration and validation datasets represent two independent samplings from the target (parent) population. The degree of difference between them should reflect the variations that can be expected associated with the future measurement situation in which the regression model is to be

used for prediction purposes [4]. It is easy to see that test set validation is the only approach which honours all these requirements, ibid.

In 2-way chemometrics there are steadfast different opinions regarding how exactly to divide the data in calibration and cross validation sets or segments [6]. From so-called full cross validation (leave one out) on the one hand, to two-segment, so-called "test set switch" on the other; the latter represent a singularly unsatisfactory choice of terminology, as there is no "test set" present at all. It is always possible to use any intermediate number of segments from the list: 2,3,4 ....N. The relationships between test set validation and these systematics of cross validation remain an area of some confusion in conventional 2-way multivariate calibration [4]. In multivariate image analysis, however, distinct and special considerations are required to which this paper is dedicated.

There are two major characteristics in image data that are rarely found in 2-way data. Most striking is the number of “objects”. In a conventional video image (~500x700 pixels), there are more than 350.000 “objects”, i.e. pixels, in the range [0..255]. Removing any single object from this amount of data is not going to change the model adequately to perform any useful validation [4]. Also, calculating 350.000 sub-models, full cross validation, is not very tempting.

Secondly, and much more important to consider, is the large *redundancy* that exists in image data. Pixels lying close together in the image space are likely to represent the same object, and therefore often have closely similar values. Two-block data sets, for example in which every second pixel, say, is to be used for validation, would necessarily produce two almost identical images, clearly leading to inferior validation, *ibid.* This would correspond to some spatial (image space) segmentation scheme. With knowledge of object selection traditions in 2-way data analysis, the reader might well alternatively ask: “Why not simply use random sampling then?” This would correspond to a notion of a fair "blind", automated segmentation strategy. Again, consider the very large amount of data (pixels) present. Sampling 50% randomly out of 350.000 objects would most likely again simply produce two practically identical datasets. The last refuge from frustration of trying to generalise from the well-known 2-way regimen into the 3-way MIA/MIR realm will probably be to throw ones hands in the air: "Then use a larger number of segments, 10 or so!" - We shall show below that all such "blind" segmentation strategies are doomed to failure in the multivariate image regimen, irrespective of the actual number of segments chosen - if not specifically related to the covariance structure in the multivariate image.

In fact, multivariate image analysis requires a complete reconsideration of relevant strategies for selecting relevant data sets for calibration and validation. A new a strategy called "*guided random sampling*" is suggested below. In guided random sampling the *user* decides how the data is to be divided into the pertinent sets. This is neither done randomly, nor by a pre-specified "blind" number of segments, but with very specific respect for the *empirical data covariance structure* present (in the score feature space). A different angle from which to attack the data segmentation problem is required. Following the MIA experience this angle is to be found in the score-space.

Nomenclature

The following notation is used:

X	Matrix of predictor variables
Y	Matrix of dependent variables
y	Y-vector
T	Matrix of X-scores
U	Matrix of Y-scores

CASE STUDIES

For illustration purposes, several examples mostly based on already published multivariate image data sets will be used [7, 8]. The master dataset consists of a 512x512x8 image, the Montmorency Forest experimental data set [7, 8], where the channel with lowest wavelength is here chosen as the **Y**-image in the present context. This is not to be understood so that we suggest to predict this channel from the remaining seven others (although this actually might be an excellent solution for recovering a "corrupted" channel, which is often enough met with in remote sensing) - rather we make good Y-use of this particular channel in order to illustrate the special image regression case of Y-total, compare [2].

In figure 2 the pertinent T1-U1 score-plot from this application is shown. The cross validation challenge is here to divide this plot in, say, two sets (segments) that *both* are equally representative of the actually covariance structure present. A simple two-split in this plot may easily give rise to a significant difference between the subsets if the data structure does not comply well with a simple joint multivariate normal distribution

assumption. In multivariate image analysis we have yet to find such simple relationships! Some objects in one set will not be equally represented (if at all) in the other, and validation may easily tend to become unbalanced. A(ny) two-split - alone - would almost always be in danger of being unbalanced.

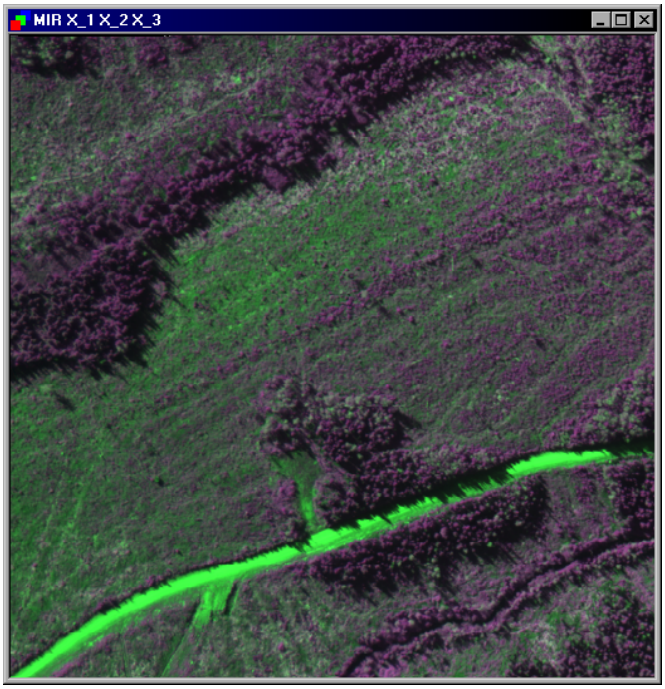


Figure 1. The scene space master image data which will be used for illustration, the Montmorency Forest data set [7,8]. The image consists of 512x512 pixels in eight bands (Channels, variables). Here it is represented by channels 1-2-3 as R-G-B.

To solve this problem, we suggest that the data set - generically - is divided in eight segments, sampling both *along* as well as *across* the dominant covariance data structures in the following way.

Initially the data is split in two halves along the main covariance direction. In figure 2 this would be a line passing through the two modes of the highest concentrations of pixels with similar score signatures, i.e. topographic “peaks”, compare [3,8] (figure 3), which are coloured red and orange in fig. 3. Each of these parts should now contain approximately 50% of the objects, and all main classes should be represented - at least the classes which go along to make up the dominating elongated covariance trend. Secondly, intersecting the first line, a new line should be drawn representing the second most important covariance direction, again as judged from the pertinent MIA

score cross-plot conventions, ibid. It is important that this second direction really corresponds to what the user perceives as the second most representative part of the overall covariance structure (more examples to be given below); thus there are no requirements for orthogonality of these two salient user-delineated covariance directions etc. This gives four segments, which each ideally should contain about 25% of the objects - barring whatever "surprises" may be in waiting in the higher-order components not captured in this first delineation. This, generally oblige, axis cross delineation is all the user has to supply in order for our new cross validation procedure to take over.

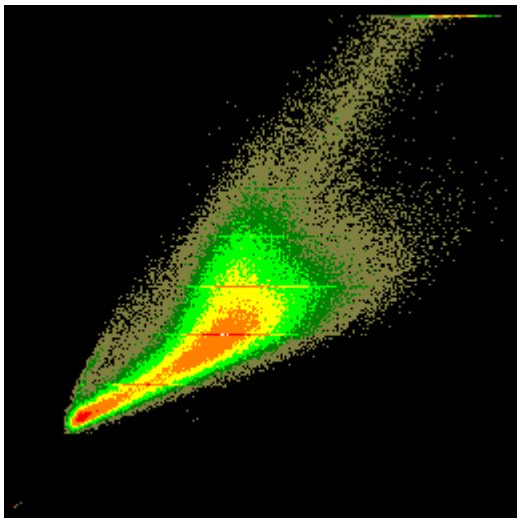


Figure 2. T1-U1 scoreplot from the MIR analysis of the image in figure 1.

After the user has drawn this second line, the software draws the four lines between the endpoints of these backbone intersecting lines. The software locates the intersection point, and finally calculates the midpoints between the corners of the outer frame. Lines are then drawn between the midpoints and the intersection point. An example of a resulting eight-segment mask is shown in figure 3. This configuration illustrates a generic eight-segment mask which it is the user's task to implement on top of a specific T-U, or T-T score plot.

With this type of mask, there are three functional combinations of subsets consisting of eight, four or two validation segments respectively. When selecting and combining sets, they should be opposite with regard to the centre point. Figure 4 shows the two compounded sets used in two-segment cross validation. In general each of these non-

overlapping two-fold division of the image covariance structure takes the form of a Maltese Cross, illustrated vividly in fig. 4.

Notice in Figs. 3 & 4 how some obvious outlying features have been excluded already in this first stage of cross validation segmentation (top right portions).

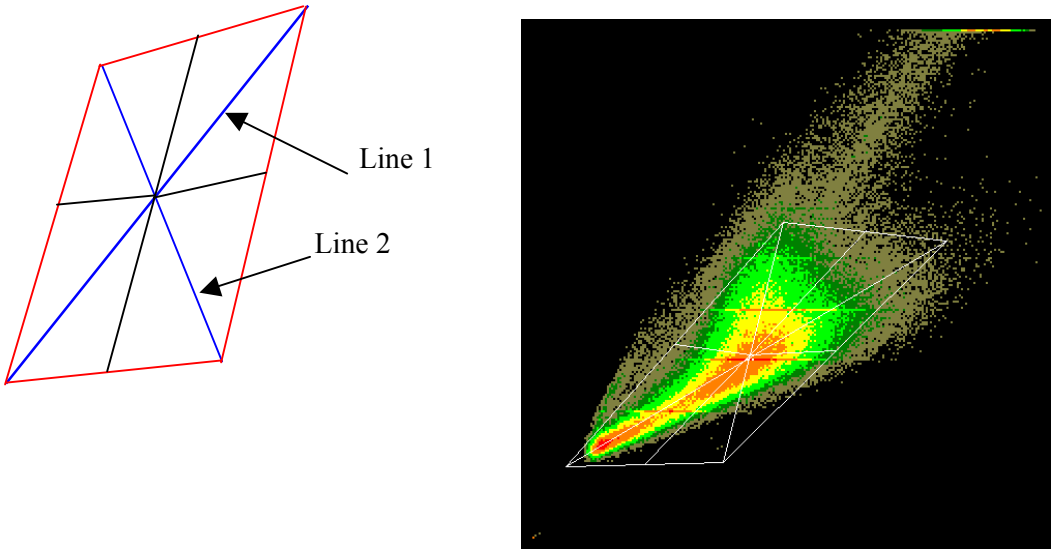


Figure 3. Left: Sketch of cross validation segment splitter initiated by two master lines drawn by the user. Right: Example of eight cross validation segments defined in a score-plot, T1-U1. Note that outlying pixels can be excluded already when delineating this mask.

Case 1: Full Y-image

The first case is a study of what was found in [2] to be a comparatively rare situation in image analysis; the full Y-image. In this situation, each object in X, each pixel, also has a corresponding representation in Y. This furnishes a particularly illustrating example of the new image analytical cross validation approach to be outlined. A more usual situation is studied in case 3.

While figure 4 shows the two validation data sets in the scoreplot, figure 5 displays the same data in image space. Pixels marked with white colour is used in the set.

Some outlying parts of the data was left out of the validation set entirely, because these pixels were identified as *outliers* already when delineating the problem-specified Maltese Cross region of interest. Alternatively this built-in outlier remover can be refined by making a local model [3,8] prior to the cross validation, allowing only the specific, problem-dependent objects of interest to be represented in the scoreplot.

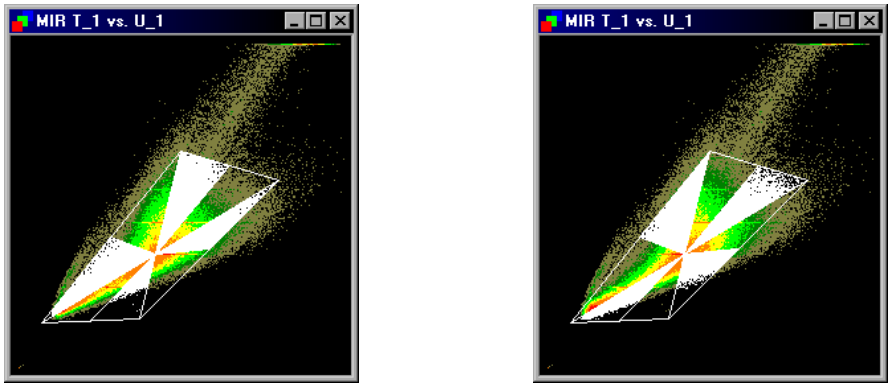


Figure 4. The two complementary "Maltese cross" validation data sets selected in the T1-U1 scoreplot shown in figure 2. Note how both achieve good data structure representation.

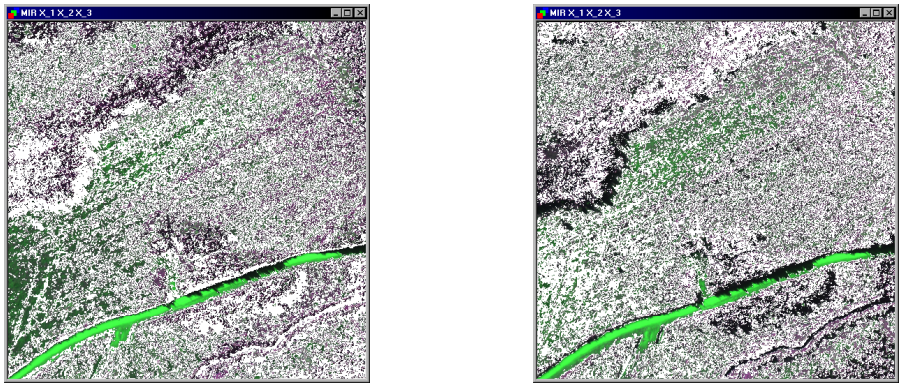


Figure 5. The two complementary validation segments selected figure 4 projected to image space. Note how both achieve satisfactory coverage and spatial representation.

Studying the images in figure 5, it should be fair to say that these two data sets represents approximately the same objects at the scale of the overall, full FOW image, with only a small difference at the most detailed levels. What you SEE in the score space rendition, fig. 4, is exactly what you GET, fig.5. The user has the full ability to iterate his or hers first tentative delineations of the Maltese Cross configuration by careful inspection of the RESULTING disposition of the two compound, non-overlapping scene space renditions, fig. 5, until a satisfactory results has been achieved.

Figures 6 and 7 shows what happens if eight segments were to used *independently as in a conventional eight-segment cross validation*. Obviously there are very great differences between these eight datasets, in fact there is an absolute certainty that these sub-models will be totally incommensurable with each other. This is a dramatic illustration also of the general cross-validation "problem" when the relationships between the X and the Y-space is more complex. In the present image analysis example, it is evident what goes wrong, were one to use an eight-segment (12.5%) cross validation scheme.

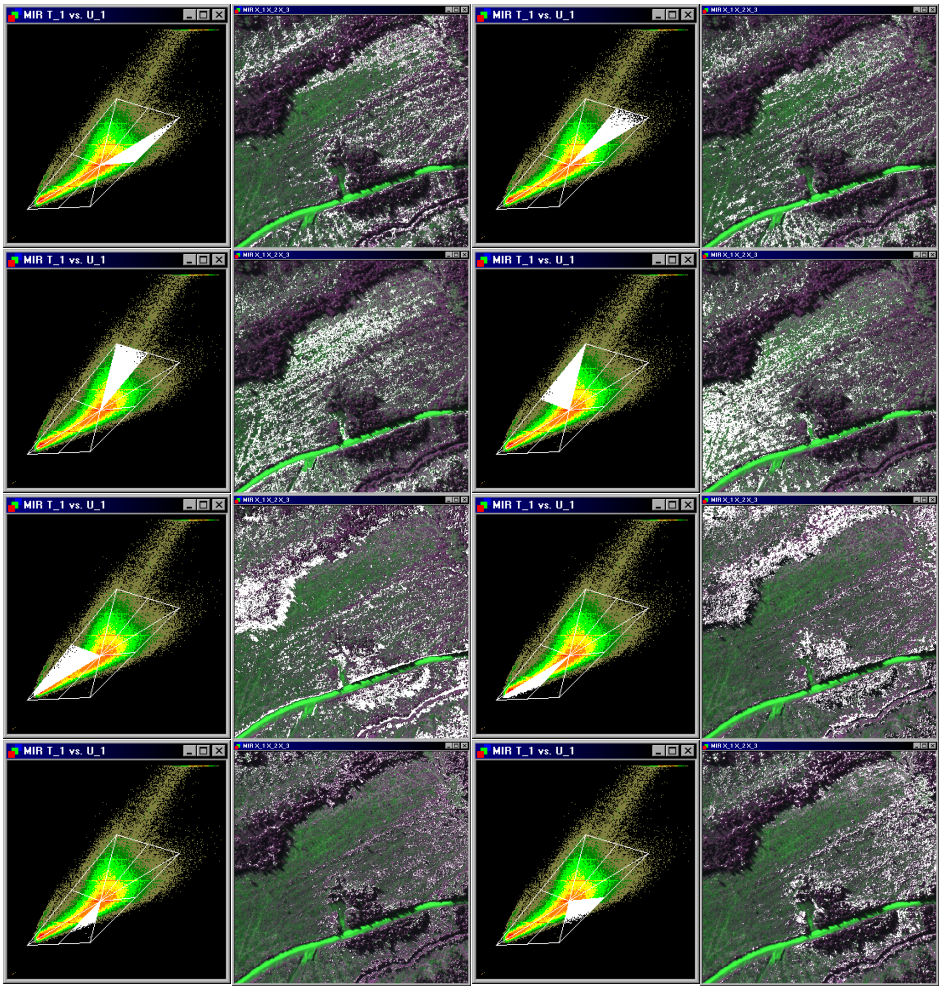


Figure 6. Eight individual validation segments in the T1-U1 scoreplot and the corresponding image space. Note how none of these achieve neither data structure nor spatial representativety.

Case 2: Problems

It is possible to run into problems with this approach if great care is not taken in the CV segmentation step however. If the segments are too small, they will very probably not be representative for the entire dataset. Another possibility, as will be shown in this case, is failure drawing the optimal guiding lines. Figure 7 and 8 shows what happens when the guiding lines split the data in a *off-centred* fashion. Clearly these two Maltese Cross configurations are NOT making up a good, balanced 50/50 cross validation bases. As can be seen even a small off-centred two-split has a dramatic effect on the two relative datasets because of the very high number of similar pixels making up the covariance backbone of the data structure. One dataset is provably very different than the other with very obvious poor, non-representative validation results to be expected. The current approach is thus very sensitive to the precision of - and the understanding behind - the user-interaction.

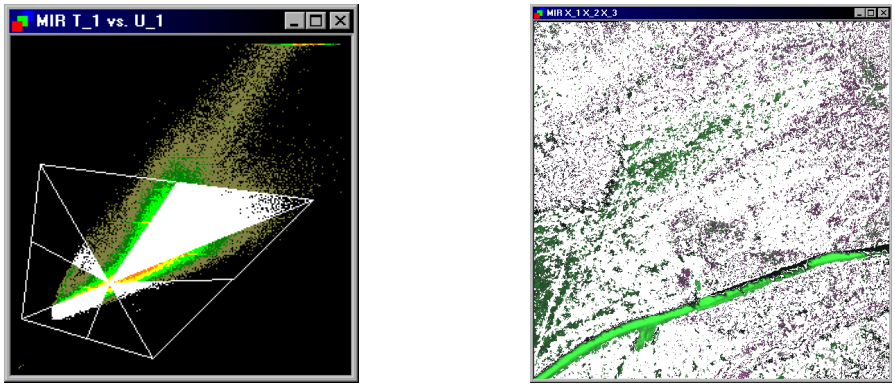
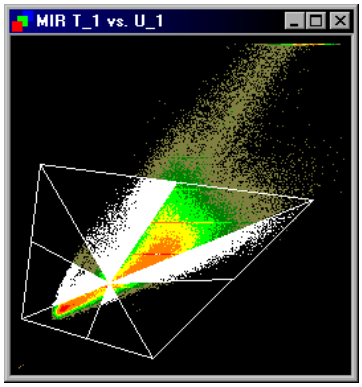


Figure 7. Corresponding scoreplot (T1-U1) and image for off-centred Maltese Cross. The complementary 50% segment is shown in figure 8.

Another potential problem is when the modes (the "peaks") in the scoreplot does not lie on a straight line. If there are more than two peaks of interest, drawing a representative two-split line through them is practically next to impossible. This problem is illustrated well by a scoreplot from a different representative data set, also from [2], illustrated in figure 9. This example illustrates with all clarity why multivariate image analytical endeavours usually are of an order-of-magnitude more complex than in the ordinary two-way regimen.



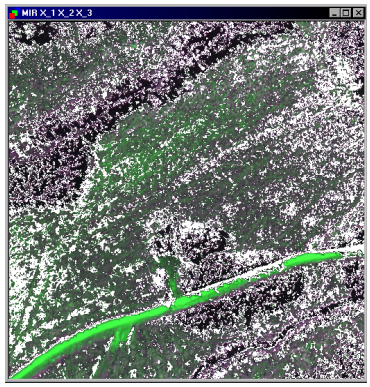


Figure 8. Corresponding scoreplot and image for off-centred Maltese cross. The contrary segment is shown in figure 7.

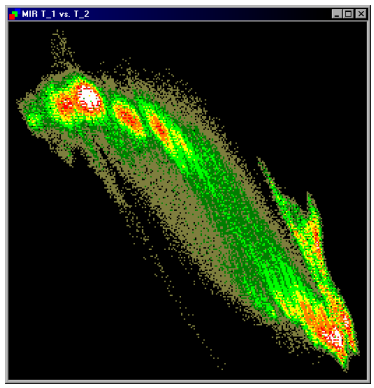


Figure 9. T1-T2 Scoreplot from a complex dataset showing a 7-8 mode ("peaked") curved data structure. Observe how it is apparently impossible to apply a Maltese Cross segmentation on a data structure as complex as this.

Case 3: Y-grid

More commonly than the full Y-image, is when X and Y are constructed as grids from several smaller images. This is a useful approach when making a reference dataset as a basis for a regression model. A typical grid image is shown in figure 11. This image consists of 6 smaller images of different sausages. The corresponding Y-image contains the overall fat-content for each sub-image. The fat content is represented as a grey-level as shown in figure 10. This data set-up was discussed extensively in [2] where used as a vehicle for explaining the concepts of MIR, Multivariate Image

Regression. In this particular case the objective is the be able to predict the average fat content in the six heterogeneous sausages (left in fig. 10).

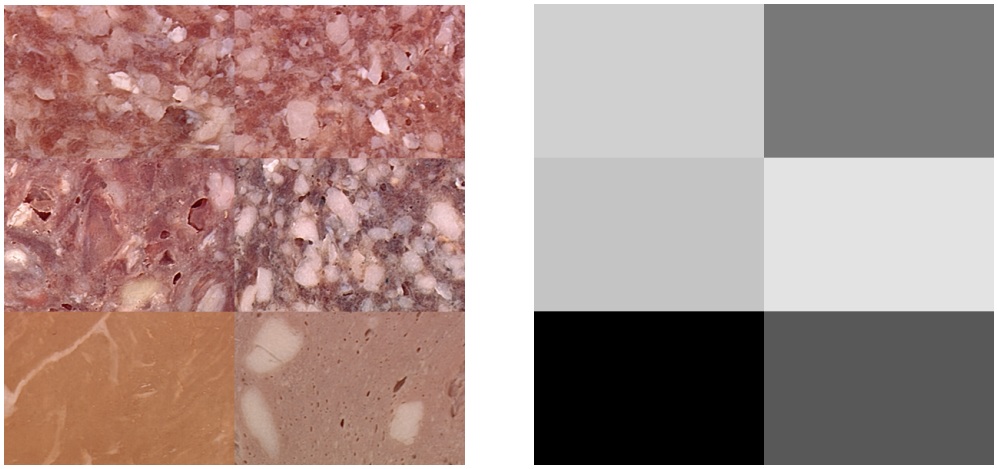


Figure 10. Illustration of the Y-grid MIR case. Six sausages, left: (X-image, variables 1, 2 and 3) and corresponding fat-content, right (Y-image).

As can be seen from the Y-image in figure 10, there is no unique Y-value for each pixel in X. This phenomenon occurs when an overall value is to be predicted from an image, and it has a somewhat negative effect on the T-U scoreplot. This effect is shown in figure 11 the pertinent T1-U1 scoreplot from the sausage data.

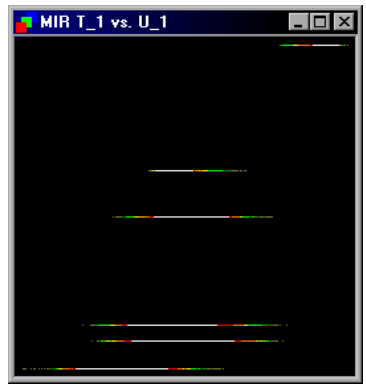


Figure 11. T1-U1 scoreplot from the sausage fat prediction case. Each line represents a specific Y-value, or sub-image, compare figure 10 (right).

In figure 12 it will be demonstrated that applying the eight segment Maltese Cross scheme in a T-U plot, as the one in figure 11 is not at all straight forward. The nature

of the T-U plot in grid cases will force an uneven distribution in the image-space, almost no matter how the eight-fold segmentation mask is delineated. Also observed how the score space delineations are very difficult to evaluate because of the extremely discrete nature of the Y-levels present in a Y-grid case; for full details, see [2].

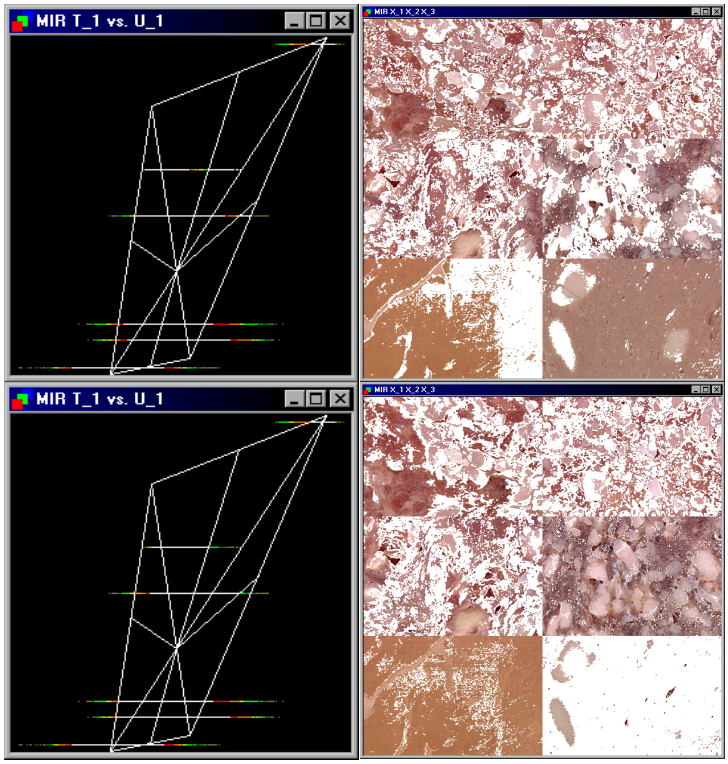


Figure 12. Y-grid case, two-block segments from selection in T1-U1 plot. Note extensive unbalance in the image space (right).

In figure 12 it is evident that especially the two lower and the middle right X sub-images are very poorly represented in the complementary validation segments. This is even more so if eight individual segments were to be used, as was shown in the first example in figure 6. To save space, this is not repeated for the current example.

Thus what seemed initially to be a good idea, i.e. the "Maltese Cross" eightfold cross validation segmentation in the TU-score space, on further inspection has proved to be at best a very sensitive approach - in fact in would be wrong to say that it has proved its reason for existence convincingly.

It can be shown, however, that this is merely a question of application. The critical point is not so much *how* the lines are drawn in the plot, it is *what plot* the lines are drawn in. So far, the procedure has been applied to plots where there are strong

correlation in the data, and physical objects have specific locations too, i.e. the familiar low-order score plot(s), e.g. T_1-U_1 etc. which all play a very dominating role in conventional 2-way multivariate calibration [4]. Chemometricians will be familiar with the fact that in the score space, the first dimensions contain the most structured parts of the data, while for the higher-order components there is bound to be less and less variance etc.

With this in mind, the next, perhaps surprising step in the present image exploration will focus explicitly on this higher-order score space.

Figure 13 shows T-scores 4 vs. 5 from the master Montmorency Forest example. What is interesting in this plot, is that most of the *structural* information is now *orthogonal* to the data delineated in this figure. This indicates that the current plot is *well* suited as a starting point for the cross validation data segmentation.

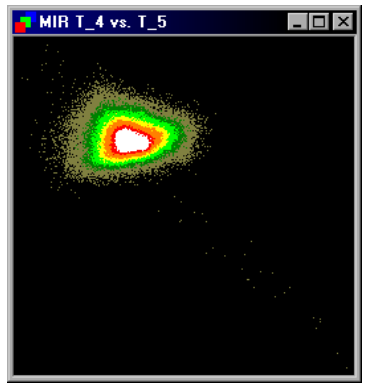


Figure 13. Alternative higher-order components scoreplot (T4 vs. T5) from the Montmorency Forest (figure 1).

Below, a Maltese Cross cross validation segmentation has been applied to the scoreplot in figure 13. Figure 14 shows the resulting two non-overlapping segments both in score space and image space. As can bee seen from the figure, there is now a very satisfactory even distribution in the two segments (and only with very close investigation, some minor differences can be found between the image-space representations though, which have to do with **shadows** mainly).

In figure 15 this is further illustrated by examining the eight segments separately. The conclusions from figs 13 - 15 are very clear: when delineating the new image analytical eight-fold cross validation segmentation in some appropriate higher-order score space rendition, in which most of the substantial data structure is orthogonal, the documented sensitivity has been controlled completely.

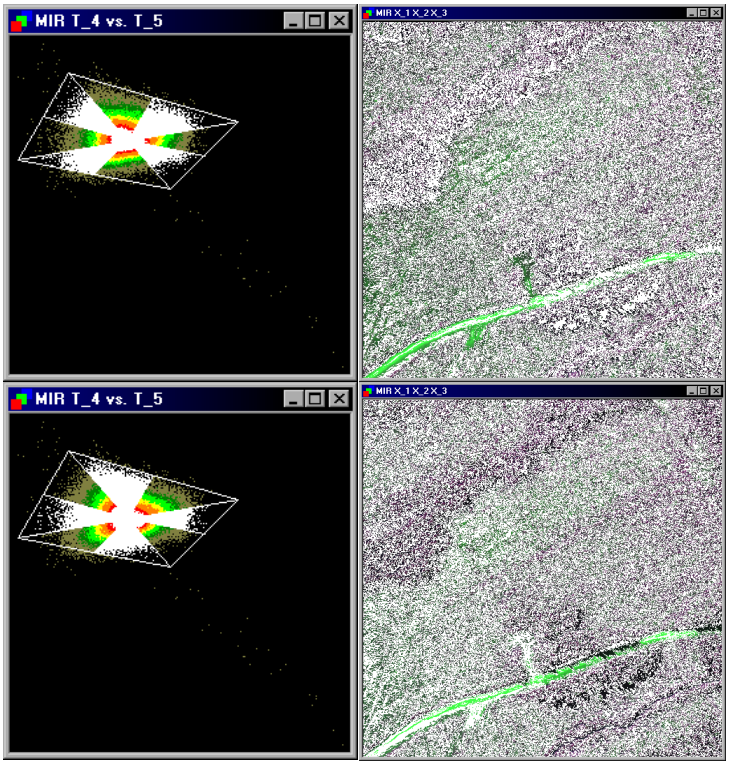


Figure 14. . Maltese Cross validation segments selected in higher-order T4-T5 scoreplot in figure 13. Note excellent data structure as well as image (spatial) coverage and representativity.

It is still evident that an "even" rotation of the segments in score-space, leads to extremely opposing unbalanced pixel divisions in the corresponding image space. From this it is necessary to conclude that many such segments must always be *combined* to form larger fractions of the entire field-of-view, e.g. two 50% segments as in figure 14.

Stepping back to the difficult Y-grid example (sausage fat-prediction), it is now interesting to see how this higher-order components approach will behave. Using T-scores 5 vs. 6 and drawing the two lines that split this data set in as equally representative fashion as possible produces the segments shown in figure 16.

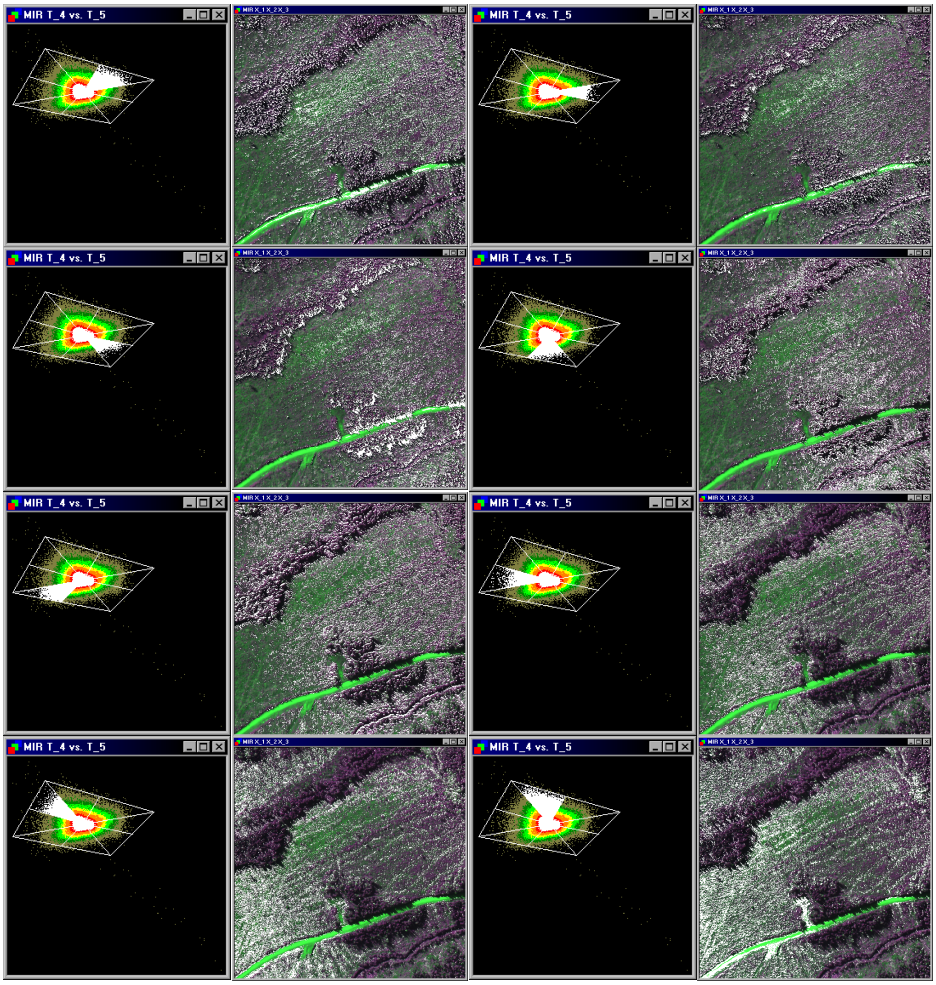


Figure 15. Eight segments from the T4-T5 scoreplot cross validation splitter shown in both score- and image space. Note that an acceptable representation has now been achieved in both score- and image space; compare figure 6.

Compared with figure 12, figure 16 now shows a strikingly more uniform distribution of one validation segment in the image with respect to the complementary calibration set - and there are only a few, minor differences. Overall, this partition should lead to a realistic validation of the prediction model performance even for this very complex difficult data structure.

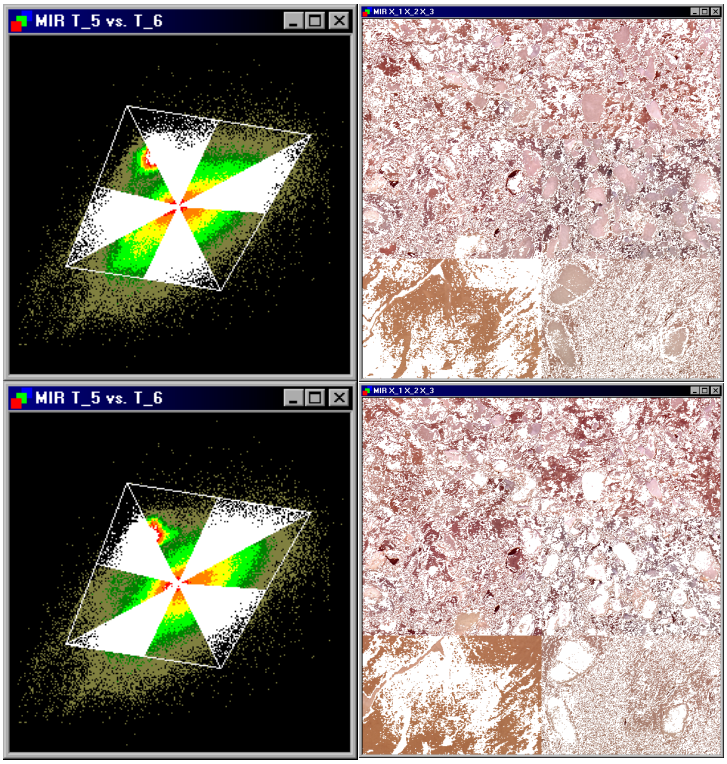


Figure 16. Cross validation segments in [T5T6] score- and[1,2,3] image space for the sausage fat prediction example. Note that an acceptable representation has now been achieved in both score- and image space. Compare figure 12.

Case 4: Cutting to the bone

One of the key features in image analysis, mentioned in the introduction, is the huge redundancy in this type of data. Having 350.000 objects describing, say typically, 10-20 classes is obviously an overkill. In MIR-cases where reducing this redundancy is essential, it is possible to reduce the number of objects dramatically by a simple procedure, compare also [8] in which this case was described for MIA. The suggestion is shown in figure 17 in the form of the curved (hand-drawn) line, where the number of objects have been reduced to a small fraction of the original, but deliberately covering all the important classes of interest in the image. This is so because it has been drawn specifically to "cover" the most dominating global covariance trend of the image feature space. Since this mask is positioned directly along the "topographic" highs, compare [8] for full details, it will - per force - be maximally representative for the essential data structure present while at the same time allowing for the exclusion of all similar pixels lying outside its width (typically 1-3 pixels wide) without any risk of

loosing out on the most representative pixels. Observe how we have made use of this feature in the so-called "pred-meas" plot (predicted vs. measured), well-known from conventional 2-way multivariate regression validation. Thus for fig. 17 below:

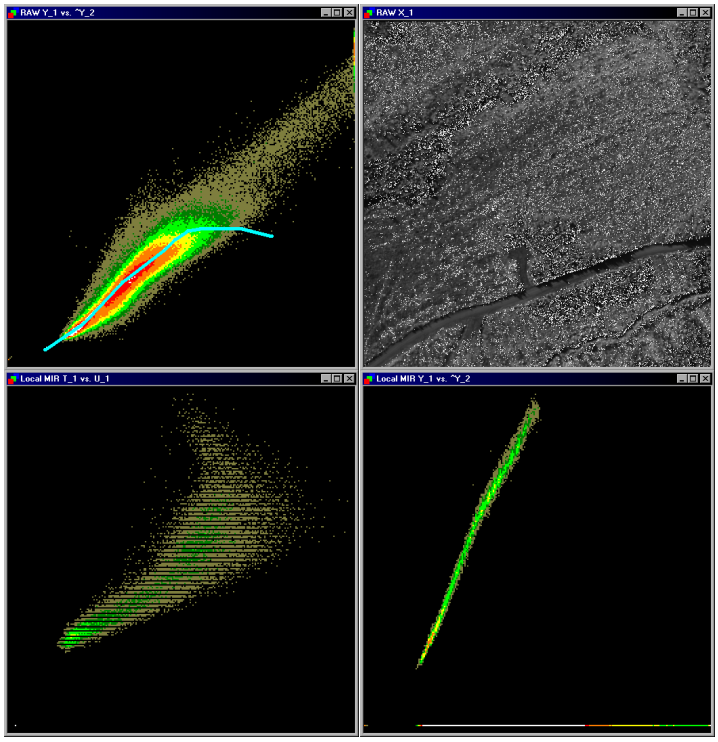


Figure 17. Freehand line covering the essential covariance structure in a Pred.-Meas. plot. After a local MIR model has been created on this basis, the corresponding T1U1 (lower left) and the Pred.-Meas. (lower right) plots are shown, validating this type of representative sampling of MIA/MIR data.

Starting out in the strongly correlated pred-meas plot, a one-pixel-wide line is drawn covering the main data of interest. This line emulates the global covariance trend as best as at all possible. All objects (pixels) covered by this line only, are then used as objects in a new, *local model* [3,8]. This model will contain far less objects, and the redundancy in the data will be strongly reduced. In figure 17, a T1-U1 plot is shown at the lower left. This can now be used as a starting point for the cross validation segmentation. The corresponding *local model* pread.-meas plot is shown at the right in the figure.

Some comments are required for the last figure. The points and line that can be observed in the lower part of the plots, represent the objects that have been **left out of the model**. In the calibration procedure, they have been removed from the data modelling, but for image

displaying purposes, it is necessary to include these pixels. To avoid them from interfering with the image, they are set to zero-value, and are displayed black in the image. The lower-left point is hence the (0,0) coordinate, as all score-values are scaled in the range [0..255] to optimise their display.

DISCUSSION AND CONCLUSIONS

We have shown that the new approach in which segmentation is done based on the *orthogonal* data representation in *higher-order* score components, is of a powerful and general nature, which in most cases will enable a realistic two-split cross validation (approximately 50/50). Segmentation following this approach takes the form of two non-overlapping "mirror" Maltese Cross configurations, each made up of four "arms". The Maltese Cross is designed specifically to allow equal (but non-overlapping) neighbouring segments in parallel along both the user-defined axes of the mask (figure 3). This enables a near-optimal representative split of the training data set across all covariance structure directions, precisely because of this *compound* nature.

We have also shown that considerable care is needed when employing this feature on the alternative lower-order component plots available (e.g. T1-U1), in which a rather large "off-centre" sensitive was demonstrated.

In general it is not recommended to use cross validation in multivariate image analysis with a number of segments higher than two, and then *only* in the form of the Maltese Cross (sic) - due to the much higher complexity of the covariance structures for this type of data relative to the experiences from the conventional 2-way realm.

In multivariate image analysis, there is usually a high degree of redundancy in the data. In such cases with relatively few physical objects (classes), data reduction with local modelling should be considered prior to validation. We have delineated a simple approach for this – the one-pixel-wide swath across the backbone of the dominating data covariance structure(s).

REFERENCES

1 Esbensen K., Geladi P. & Grahn H. 1992: Strategies for multivariate image regression (MIR). Chemometrics and Intelligent Laboratory Systems vol 14, pp. 67-86

2 Lied T.T. & Esbensen K. 200?: Principles of MIR, Multivariate Image Regression -I: Regression typology and representative application studies. In preparation.

3 Geladi P. & Grahn H. 1996: Multivariate image analysis. (Chichester: John Wiley & Sons) ISBN 0-471-93001-6

4 Esbensen K. 2000: Multivariate Analysis in Practice, 4th edition. CAMO ASA. ISBN 82-993330-2-4

5 Lindgren, F. 1994: Third Generation PLS. PhD thesis, Umeå University. ISBN 91-7174-911-X

6 Martens H. and Næs, T. 1989: Multivariate Calibration. Wiley & Sons . ISBN 0-471-93047-4.

7 Esbensen K., Edwards, G. & Eldridge N.R. 1993: Multivariate Image Analysis in forestry applications involving high resolution airborne imagery. 8th Scandinavian Conference on Image Analysis- SCIA'93, pp. 953-963

8 Esbensen K., Lied T.T., Lowell K. and Edwards G. 200?. Principles of Multivariate Image Analysis (MIA) in remote sensing, technology and industry. In preparation.

Paper V

Image-Analytical Quantitative Monitoring of Heterogeneous Mixture Processes: Angle Measure Technique (AMT) vs. Multivariate Image Regression (MIR)

Thorbjørn Tønnesen Lied, Inger Hedvig Matveyev,
Dawn Angela Field Karlsrud, Jun Huang and Kim H. Esbensen*

Applied Chemometrics Research Group (ACRG), Department of Technology (TF)
Telemark University College (HiT), N-3918 Porsgrunn, Norway

* Corresponding author. E-mail: kim.esbensen@hit.no

CONTENTS

Abstract 1

Introduction 1

AMT (Angle Measure Technique) 4

The MIR Concept 8

Analysing the primary MIR prediction results 10

Thresholding 10

The Mean Grey-level Value 10

Histogram Calibration - Extended MIR (MIR^{+}) 10

Data Presentation 11

Two-Component dry Granular Mixtures 12

Three-Component Granular Mixtures 14

On-line Minced Meat Mixing Fraction Specification Control 15

Imaging system 17

Results 19

Two-component Granular Mixtures 19

Black Pepper and White PVC 19

Coriander & White Pepper 20

Grey and White PVC-pellets 21

Grey PVC-pellets and green beans 22

Three-Component Granular Mixture 23

PLS1 (y_1 : Peas mixing fraction) 23

PLS1 (y_1 : Maize mixing fraction) 24

PLS1 (y_1 : Carrot mixing fraction) 24

Minced Meat 25

Bovine 25
Pork 26
Bovine & Pork 27
Discussion 27
Two-component mixtures - overview of results: 27
Three-component vegetable mixtures: 28
Minced meat mixtures: 28
AMT or MIR^{+}: 28
Acknowledgements 35
References 36

ABSTRACT

Selected two- and three-component mixtures are studied by image analysis plus *chemometric data analysis*, specifically AMT (Angle Measure Technique), MIR (Multivariate Image Analysis - and the recent extension termed MIR^{+}) and PLS-R (Partial Least Squares Regression). The present studies comprise a first foray regarding the possibilities of continuous mixing process - and product monitoring (homogeneity, on-line mixing fraction quantification etc.) using image analysis as the primary data capture facility.

We study three very different types of mixing systems, i.e. dry two-component powder systems, frozen three-component vegetable mix systems and a minced meat mixing system - the latter two of which constitute real-world industrial systems of current economical interest.

Results show that the present line-up of chemometric image analysis and data analysis methods are fully sufficient to outline a framework for automated process monitoring systems. The two-component systems are also representative of a much larger study (barely initiated) on the possibilities of predicting the ultimate propensity of mixing systems, based only on standard image analysis characterisation *plus* the necessary chemometric data analysis.

We conclude that *both* the AMT and the MIR^{+} approaches are suitable for the realistic tasks specified in the current studies, both with satisfactory relative prediction accuracies and precisions as estimated from comparable cross-validations.

INTRODUCTION

Mixing, blending, homogenisation of granular poly-component materials is of great importance in modern manufacturing and in large-scale process industries. These processes have been studied intensively within the field of powder science and technology. Despite it's importance, however, a full understanding of granular mixing processes is surprisingly limited. Often quite unexpected segregation can occur, even in what was thought to be well-designed mixers or blenders, for example when batches are mixed just slightly "to fast" or for just a trifle "too long" etc. A recent overview highlighted these difficulties graphically with great impact[1]. Here was outlined in detail the many factors involved in determining the ultimate outcome of a particulate matter mixing. We have determined to begin a major effort of mapping the interplay between these instrumental factors - only based on a direct image analysis characterisation of the components involved plus whatever *a posteriori* intricate (or

simple) data analysis needed. I.e. we want to find out the possibilities - and limitations - of being able to use only a standard, non-invasive image analysis recording technique of the *in-situ* appearances of the end-member components involved - in order to be able *to predict* the final outcome of physical mixing- and blending testing under real-world conditions (this latter to be carried out in realistic full-scaled certified mixing experiments of the type reported in [1]). For this predictive effort, as well as for the presumably rather complex post-imaging data treatment necessary, we turn to *chemometrics*, which has been used in connection with powder science and technology only in the last five years, but with rather spectacular results [2, 3, 4 , 5, 6].

As a first side-benefit of the above major experimental research program it was found that a selected few of the many types of poly-component mixtures involved, also could serve a more limited purpose of illustrating new, modern and efficient possibilities for on-line mixing process characterisation, together with a few different, but closely related, types of mixtures, all of which have very great significance as general representatives of industrial systems in need of reliable, precise and accurate process monitoring (mixing process monitoring). On the market today there are to be found many types of such monitors to be sure, but they are all more-or-less rather dedicated systems directed towards rather narrow classes of materials, powders etc.

We want to develop a completely *generic* image analysis_cum_chemometrics system, to be based on existing, inexpensive off-the-shelf digital video camera technologies only. Thus the "new" elements in the system we are developing will mainly be the problem-specific chemometric image data analysis (AMT, MIR^{+}) and related quantitative prediction facilities involved (PLS). We shall use very simple digital video imaging data capture in our studies, which never-the-less is of exact industrial standards.

Specifically the present work attempts to predict quantitative mixing fractions on a selected set of (very) different mixing series. We *simulate* on-line monitoring of representative mixing processes, by preparing precise (v/v %) quantitative fractions of the granular materials involved and subjecting them to the above camera under direct industrial process monitoring conditions.

Of the methods employed here, the Multivariate Image Regression (MIR) approach is used to establish a relationship between video imagery data (X) and functional granular properties (Y) (here we limit ourselves to addressing quantification of the homogeneity of mixings, but in the major research program mentioned above, we shall also address

a range of other, standard powder characterising functional properties). Images taken of granular materials are bound to contain inherent information concerning the geometric shape(s) of the individual particles, their sizes, or size distributions, surface roughness, irregularities, smoothness, etc. This type of imagery will also contain information related to the ensemble of particles, to the particulate aggregate, or powder etc. as it were. Both these basic types of characteristics can be related to the bulk granular functional properties and mixing fractions, using proper *multivariate calibration* (PLS-regression), based on derived AMT-spectra[2] and/or MIR analysis techniques [8-9]. The major methodological objective of the present work will be to *compare* these fundamental two (the AMT-based and the MIR-based) prediction possibilities.

Two laboratory systems are studied below, a selected suite of four dry two-component powder systems (representing very varying colour -, reflectance -, as well as morphological contrast) and a system consisting of vegetable mixes of three end-components. The first set is used to study the more fundamental factors governing the mixing processes and - results, while the second set is a *bona fide* industrial mixing process monitoring example.

In addition to this study of granular mixtures, a real-world food mixing process involving minced meat (mincing meat and fat at industrial scales) is examined, using a newly developed industrial standard mixer ("IDE-CON"). To be able to satisfy their customers, producers of minced meat products are critically dependent upon reliable, essentially real-time on-line measurements of fat content in their products, as only very small deviations from the health authority specifications are accepted. Traditional off-line fat measurements are time-consuming and certainly not continuous. Image analytical on-line measurements would be very preferable, if feasible, also because of the possibility of actually designing systems with an ability - in principle - to be *totally representative*, i.e. image analysis solutions have the prospect of being able to inspect the *entire* production output.

A few examples of *sample preparation problems* occurred when certain sample types were introduced to the camera. Homogeneous mixtures would sometime *segregate* slightly when being poured onto the sample presentation plate in front of the camera etc. Flow - and transportation segregation is often a well-known problem in the handling of particulate matter. We made serious efforts to curtail this heterogeneity-increasing factor in the studies presented below, by *standardising* the specific sample presentation process (presenting the sample to the camera) throughout. On the other

hand what (little?) *pouring segregation* variance remaining was accepted as it indeed mimics correctly the identical problem facing industrial inspection systems. This certainly introduced more realistic variation in the analysis results obtained over that of otherwise "idealised" laboratory systems.

AMT (Angle Measure Technique)

The AMT transform, as a new *signal analysis* method, has shown potential in many areas of science and technology since its debut in 1994. It characterises the scale-dependent complexity of data such as time series, spatial data series, indeed any generic *measurement series*, in a new domain - the *scale domain*. Applications include image analysis, signal analysis, spectroscopy, analysis of drilling well log data, measurement runs in quality control, etc [2-6].

AMT has been applied on powder imagery in connection with multivariate calibration in a series of recent studies resulting from our chemometric collaborations in powder science and technology [2-6]. AMT has shown a significant positive propensity as a salient *pre-processing* facility for quantifying the textural characteristics of images. When derived AMT-spectra (see further below) are subjected to *multivariate calibration*, e.g. in the form of PLS-regression modelling, a combined facility termed MAR (Multivariate AMT Regression) has been shown to have a very wide applicability. It is especially the combined facility of being able to quantify texture features for both individual particles as well as for their aggregate (powder/mixture) characteristics which comes to the fore in these applications, allowing for materials characteristication *simultaneously over all particulate scales*.

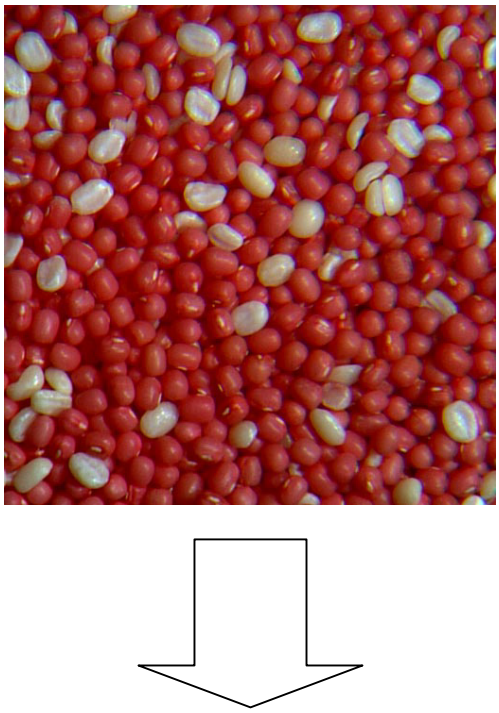


Figure 1. A representative image of a mixture of particulate matter, to be unfolded *and subjected to AMT-characterisation*

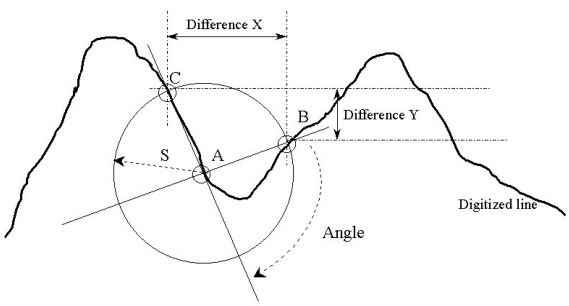
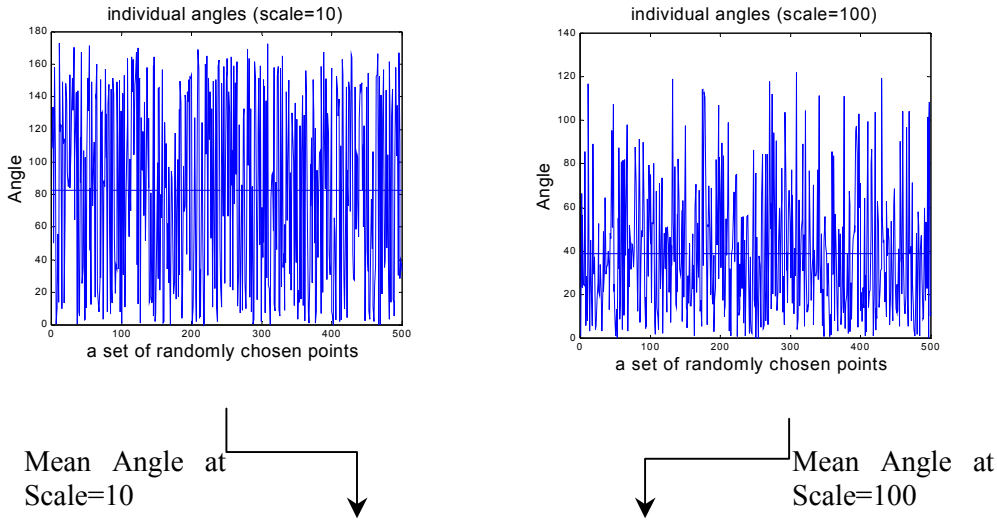


Figure 2. AMT-derivation of the so-called MA angle measure (Mean Angle). The extensive reference litterature explains parallel derivation of the MDY-measure as well [2-6].



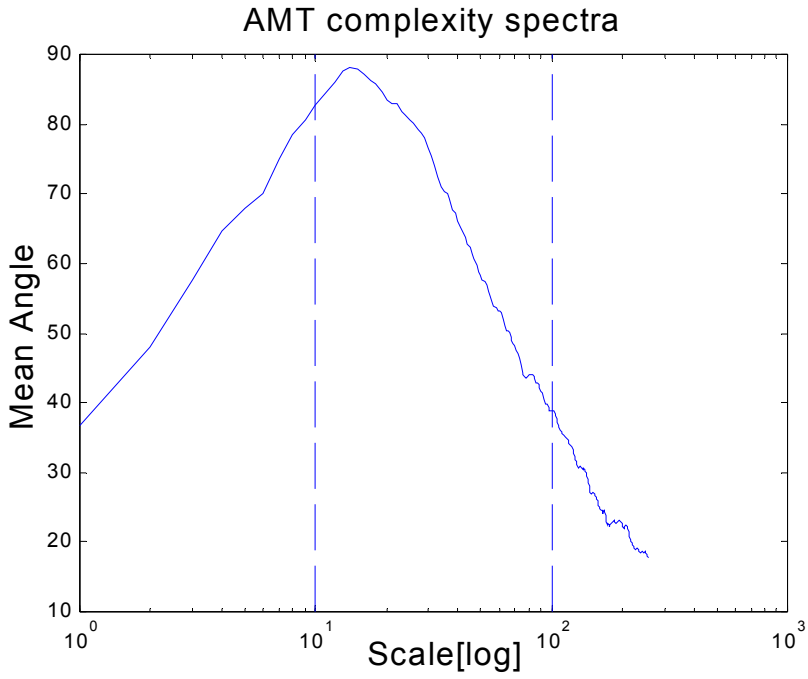


Figure 3. Illustration of representative AMT complexity-spectrum derived from a mixture image. Observe how the MA-spectrum is calculated as the mean of all individual angle measures at all scales, two of which two have been highlighted (scales = 10, 100). The horizontal axis represents "log S". MA displays a complexity "peak" corresponding to a scale of approx. 10-30. When several AMT-spectra are collected into a common X-matrix, the (log S) scale is used as the variable dimension. For full details of AMT [2-6].

The most useful aspect of the AMT transform is that the compound (**MA**, **MDY**) spectrum can be used as 1-D object vectors in multivariate data modelling (e.g. PCA or PLS). For 2-D image objects it is the *local texture* of the field-of-view which is transformed into a corresponding 1-D *linear complexity spectrum*. These complexity spectra, implicitly carry a remarkable information richness related to all scale(s).

Multivariate AMT regression (MAR) has brought a new approach to extracting information for prediction purposes from "measurement series" (of any kind), which in the present context consist of *unfolded*[†] isotropic digital images. This approach converts texturally isotropic images into 1-D multivariate AMT-spectra without loss of fidelity. It views an image in a mathematically transformed way instead of by direct visualization. The present work deals with granular powder and food particular materials as well as minced meat imagery, but applies equally well in many other similar situations.

Figure 4 below shows a schematic overview of the processes involved when using AMT-spectra of images for multivariate calibration. The MAR approach requires several steps when used for quantification of heterogeneous mixtures. Regression models must either be created based on images of *pure* mixing-components (classes), or based on a series of "spiked" concentrations of one, or more of the end-member fractions etc. The appropriate AMT-spectra are combined in a training data set related to the multivariate calibration PLS-modelling. If possible a relevant test set should also be prepared etc [7].

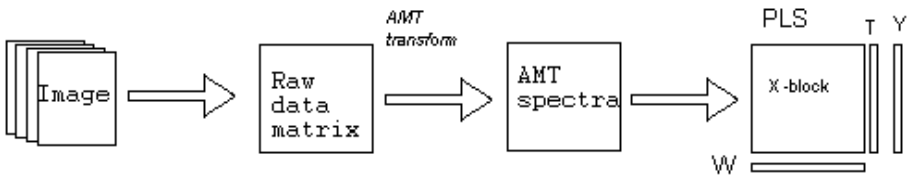


Figure 4. Schematic overview of image processing with AMT before regression calibration, MAR (Multivariate AMT Regression).

[†] *Unfolded* is here used to describe the operation of rearranging each image-channel from a 2D matrix to a long 1D vector.

For a full description of the AMT approach, see [2-6]. Suffice here to emphasise that the AMT image processing approach deals with characterisation of *contrasts* in both colour, reflectance, texture, individual grain forms and more. Thus it is not only the - perhaps more conventional - geometrical texture interpretations which are codified. Since derived AMT-spectra of an imaged material represent a *unique* scale-domain complexity/texture pattern of the image, they are well suited for calibration of images where *changes in overall texture* is an issue. Consider e.g. a series of fractions of two -, three - (or poly-) component mixtures (used as a calibration data set for the present studies). It will be seen that it is the *totality* of all these potential texture features which is changed when the mixing fractions of one (or more) of the end-members are changed. This shall be amply illustrated below. It bears to observe that it is not strictly necessary to be able to understand in all details which of the *individual* different contrasting texture factors are involved - nor *how*, or their supposed much more complex potential *interactions*. We have shown in several of the precursor investigations upon which the present work is firmly established [2-6], that the compound AMT-spectra in a sense *automatically* codifies all relevant factors and that it is the subsequent PLS-regression multivariate calibration which is responsible for extraction of *precisely* those parts of these X-spectra which *correlates* most strongly to the chosen Y-variable, which will be the pertinent mixing-fractions in all present studies.

There is thus a well-reflected reason to expect that the AMT- approach also will be successful in quantifying the mixing-fractions involved in the present experiments, but it is an open question to what ultimate levels of accuracy and precision this will be attainable. The AMT-approach will be compared below with the Multivariable Image Regression alternative, MIR - especially in a novel, extended modification, MIR^{+}.

The MIR Concept

MIR (Multivariate Image Regression) [8, 9, 10, 11] can also be viewed as a transformation of images. In this case, the multivariate image is transformed from raw data to PLS-components [12, 13, 14] called *score-images*. MIR is aimed towards being able to predict Y-images based on a regression model [15, 16, 17]. The predicted images may often constitute the final result in themselves, but can also sometimes be used for further feature extraction in several ways.

The MIR approach requires several steps when used for quantification of heterogeneous mixtures. First of all, regression models should be created from images

of *pure* mixing-components (classes). These images are combined in a training data set, as shown in Figure 5.

In this illustration three models are created, one for each class. It is possible to use PLS2 to create one general model for all classes, but predictions will usually be improved by using separate models [7]. Appropriate reference Y-images are generated so as to maximise the grey-level intensity differences between the end members. The Y-image contains the maximum grey-level value at the image positions of the training object(s), and the minimum values in all other positions. For an unsigned, 8-bit image these values are 255 (white) and 0 (black), respectively. The graphic illustration in Figure 5 is probably much more directly telling....

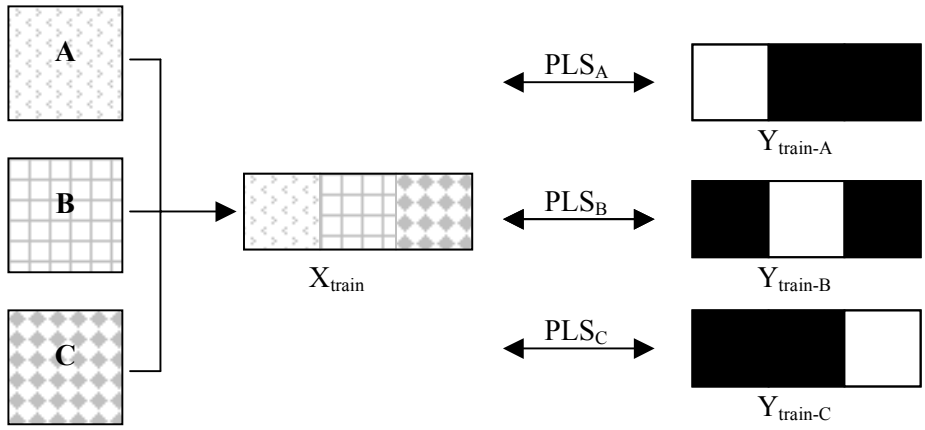


Figure 5. Illustration of MIR training set-up for quantitative characterisation of classes A, B and C of heterogeneous mixtures. A separate model is calibrated for each class, using dichotomous (white/black) reference Y-masks.

When acquiring *new images* of mixed classes - corresponding to taking an image of the product to be characterised (f. ex. on the production conveyor belt etc.), these will be Y-predicted with the models created above. If the training models have been created successfully, and if the particular input-output relation is generally linear or can be modelled by a bilinear PLS-regression model, pixels belonging to the current training class in question will be characterised by bright grey-level values ("close to white"), while all other pixels will usually be much more dark in their grey-level values. The fraction of "bright pixels", suitably defined (problem-dependent), will thus be expected to correlate to the overall mixing-fraction of the *current class* in the mixture. We have recently published several extensive MIR descriptions with a wide variety of laboratory and industrial illustrations elsewhere; see [8-17] for in-depth MIR coverage.

While AMT mainly focuses on textural and spatial information like contrasts in shape and pattern, MIR focuses on *spectral information*. This means that AMT has its preferences if texture and spatial information is important, while MIR relies much more on differences in colour etc. with which to classify and to quantify objects. It is thus expected that it is not necessarily an easy given which method to apply in a given situation. Clearly one needs a lot of experience with many, and diverse, problem-dependent data sets and applications. If both spectral and spatial information is valuable for the quantification, *combining* the results from AMT and MIR should be an advantageous possibility.

Analysing the primary MIR prediction results

Three different approaches are discussed when it comes to correlating the MIR-predicted images with the concentrations of the different mixing fractions. Two of these are univariate, and the third method is based on multivariate calibration.

Thresholding

Thresholding is perhaps the most well-known “traditional” way of analysing the frequency of bright pixels in an image. By converting the image to black and white at some critical grey-level *threshold-value* and then counting the number of "white pixels", an estimate of the concentration of the class can be calculated. If the black and white image is binary (0’s and 1’s), calculating the *mean value* will give the mixing-fraction directly.

Problems with this approach applies to noise in the data that f. ex. may result from sub-optimal lighting conditions (highlights and/or shadows), which can lead to severe misclassification etc.

The Mean Grey-level Value

Especially when dealing with two-component mixtures, in which one end-member is predicted bright and the other dark, calculating the *mean grey-level value* will correlate to the fraction of bright pixels in the image. This method does not give an answer in fractions units though, and some further (linear) adaptation of the result will be required.

Histogram Calibration - Extended MIR (MIR^+)

Instead of thresholding, or calculating a mean grey-level value for the image, the entire grey-level histogram[18] of the predicted image can be used for multivariate calibration

using PLS. This approach is the one most similar to the AMT approach, which also depends on multivariate calibration of (complexity) spectra. Figure 6 shows how MIR and 2-way PLS relates to the multivariate image **X**, the predicted $\hat{\mathbf{Y}}$-image and its histogram. The MIR model being used for $\hat{\mathbf{Y}}$-prediction has been established earlier using the approach outlined in Figure 5. The PLS-model used to predict the final mixing fractions has been established on the basis of a calibration set of several histograms with known Y-values. This is in fact a standard 2-way PLS multivariate calibration, in which the initial MIR Y-image prediction_cum_histrogram derivation can be viewed as an image pre-processing step.

This new *compound* MIR/Y-pred/histogram/PLS-approach is termed the ***extended*** MIR: MIR^{+}.

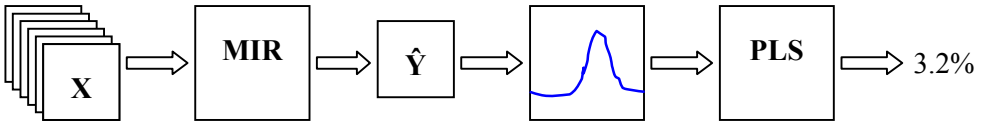


*Figure 6. The MIR^{+} $\hat{\mathbf{Y}}$-histogram prediction approach. A $\hat{\mathbf{Y}}$-image is predicted from the Multivariate Image **X** using an existing MIR model. The histogram of the $\hat{\mathbf{Y}}$-image grey-levels is used for training a traditional 2-way PLS-R model.*

It might perhaps be argued that instead of predicting an image prior to mixing fraction calibration, why not just use the raw image grey-level histograms directly? In some simple cases this *may* indeed be possible, but certainly not as a general approach. When working with a large number of video channels, it can be seen generally to be difficult to isolate just one *singular* channel that optimally enhances a single class with respect of all other classes, compare above. The powerful data compression that lies in MIR (PLS), and the extended MIR^{+}, is most often much more effective for extracting the kind of information needed for the calibration of a specific class or for quantitative mixing fraction prediction.

DATA PRESENTATION

AMT and MIR has been applied in parallel to all the data sets presented briefly in the introduction above. The data sets are further presented in a sufficient detail below - with representative imagery and accompanying explanations - in order to be able to serve as the common framework for *comparing* the alternative AMT and MIR^{+} quantifications below.

Three different mixing product types were used: two-component dry granular mixtures, three-component "wet" mixture and three different series involving mixing fat into minced meat product types.

Two-Component dry Granular Mixtures

Four representative combinations of two-component granular mixtures was selected for the present purpose - out of a significantly larger *experimental design* of nine mixing series, which have been designed to span a maximum coverage w.r.t. the three principal design variables: *colour* contrast; *reflectance* contrast and *morphological* contrast. (These design variables represent critical material factors involved in image analytical imagery, representing the primary image quality response(s) to the illumination conditions etc.). This background study specifically only addresses the AMT-prediction feasibility studies.

The four series chosen here represent both "easily AMT-modelled" systems, as well as their *distinct counterparts*, i.e. systems which did not lend themselves to fair AMT-modelling - perhaps potential candidates for the alternative MIR^{+} approach?

The first dry two-component mixture consists of whole-grain black pepper and white PVC-pellets, generally of similar grain size, Figure 7. This mixture was chosen because of its marked high spectral - (large colour difference) and textural contrasts (relatively large difference in grain form definitions). The data set contains 11 principal mixing fractions, all imaged with *four replicates*. For one component the fraction was [0%, 10%, 20%, …, 100%], with complementary fractions for the other.

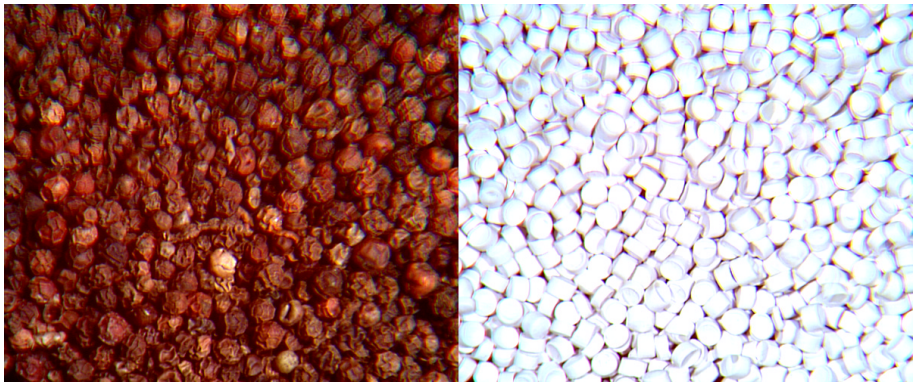


Figure 7. Whole-grain black pepper and white PVC-pellets. Training set-up of pure classes.

The second dry granular mixture involved a *whole-grain* coriander and *ditto* white pepper mixture, illustrated below. This system was chosen as a "maximally difficult

system". Thus a mixing system was deliberately *designed* to have maximally low spectral and textural, as well as morphological contrast for both end-members, i.e. In Figure 8 one observes the dramatic difference to the system in Figure 7. In this system it is decidedly not easy to distinguish between the individual grains from either pure end-member.

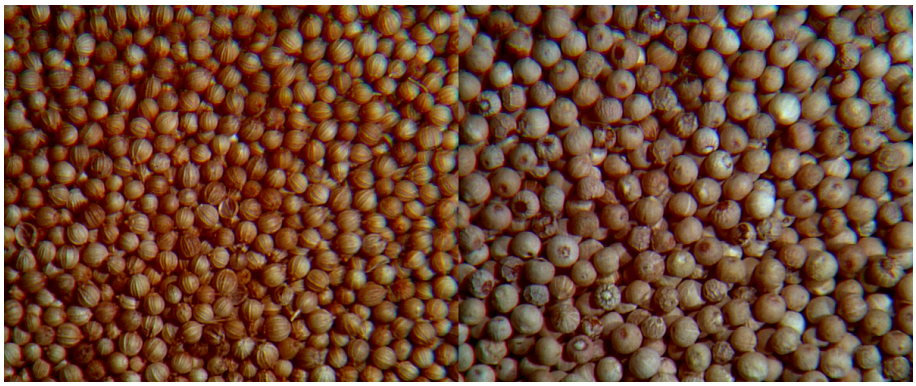


Figure 8. Whole-grain coriander (left) and ditto white pepper (right). Observe the dramatically smaller contrast compared with Figure 7.

The additional two dry mixture systems were chosen so as to represent more intermediate contrast ranges for the three design factors.

The third two-component mixture thus concerned grey and white PVC-pellets, Figure 9 with relatively high spectral contrast, but distinctly low textural contrast. There are however some important differences w.r.t. the individual grain shapes, but their average grain sizes are almost equal. Compared with Figure 7 & Figure 8, this system is clearly intermediary, as is it's close companion, shown in Figure 10 (the same grey PVC-pellets, but now mixed in with green beans).

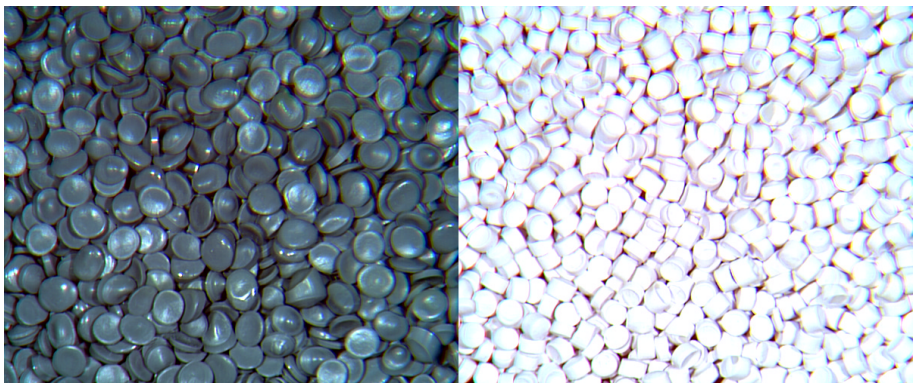


Figure 9. Grey and white PVC-pellets (of different grain shapes)

The last two-component mixture (grey PVC-pellets/green beans), Figure 10, was chosen primarily for the marked (large) contrast w.r.t. to the two individual grain shapes involved - every other contrast being in the intermediary/low range.

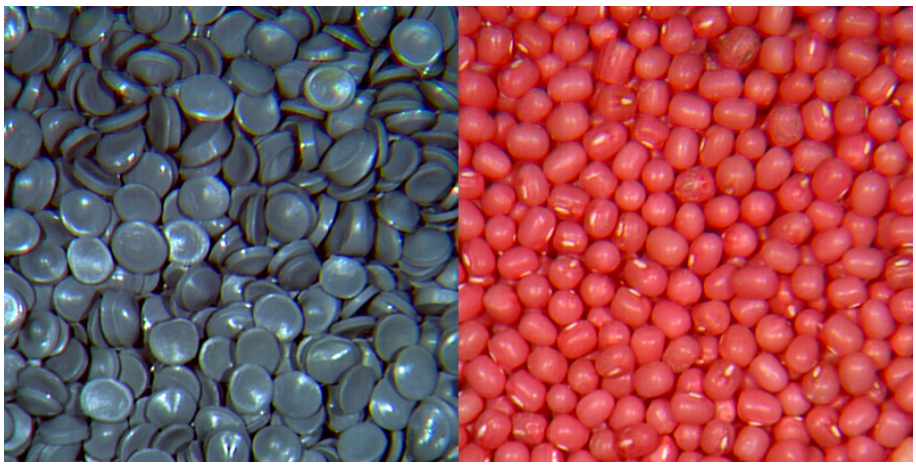


Figure 10. Grey PVC-pellets (left) and green beans (right). The NIR-Red-Green camera used (SILVACAM) is responsible for the false colour appearance of the "green" beans in this rendition. Note especially high grain shape/form contrast.

Three-Component Granular Mixtures

One major three-component example was selected for this study, a real-world industrial mixing problem directly from the production line of a Norwegian producer of vegetable mixes - in the present example we focused on an *evergreen* mix: green peas/maize/carrot (cubes), Figure 11. The producer is concerned with on-line quality control (just precise enough), which translates into an image analysis system, which should be able to predict mixing fractions of two of the three components with a precision of 8% (rel.) or better. In the laboratory experimental design used here, the fractions of each component was varied in increments 0%, 25%, 33%, 50%, 67% 75% or 100%. The imaging system is presented below, Figure 17.

For this three-component system we experienced occasional rather severe homogenisation -, and especially critical *pouring segregation* when presenting the mixed samples to the camera field-of-view, compare above ("Introduction"). It was necessary to instigate a detailed sample handling and presentation protocol, to be very strictly adhered to for all samples involved - and still some measure of residual individual sample preparation variance could be observed. This we decided to keep as it was however, for reasons of compliance with realistic on-line sample preparation in the industrial realm.

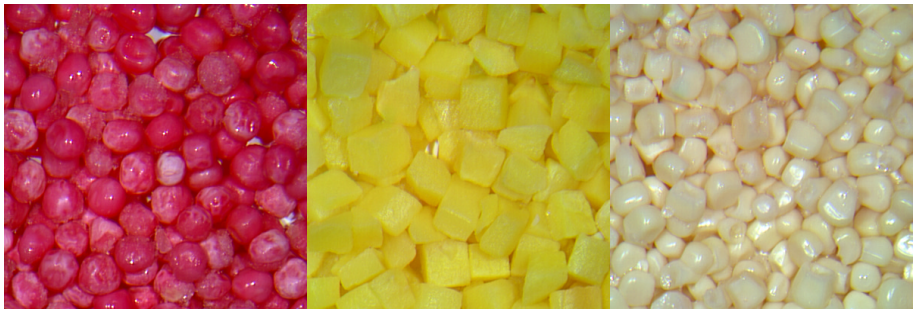


Figure 11. Green peas (left), carrot (middle) and maize (right) pure training classes(100%) respectively . SILVACAM's false colours figures prominently here, but are of no consequence for the spectral contrasts involved.

On-line Minced Meat Mixing Fraction Specification Control

Applied Chemometric Research Group is presently involved in a long-term campaign of particulate matter and powder application studies. One recent new avenue concerns outlet quality control from a novel industrial mixer (the "IDE-CON" mixing concept), which is briefly presented in Figure 12.

Figure 12.The IDE-CON continuous mixer. Note the two counter-rotating shovels of the new, proprietary IDE-CON design.

The IDE-CON continuous mixer is presently used extensively in selected test industrial sectors, amongst which mentioning of the following high-precision target examples should suffice to illustrate the importance a reliable mixer-outlet product specification

verification: on-site road tarmac mixing/blending from three raw materials (all at elevated temperatures of about 85°C; several poly-component health product manufacturing, with up to 10 components in concentrations ranging from, say 1000 ppm to typical filler status (50-90%).

The IDE-CON mixer has the added versatility of being able of *continuous* on-line *adjustment* of the blending/mixing regimens as needed. Therefore it is also used in industrial sectors and branches, for example the food - and feed producers etc. where non-invasive (indeed sterile), strongly regulated, precise control facilities are required by the authorities. As an example, from a leading Norwegian agricultural producer, our last example is related to industrial production of minced meat products - in which quick adjustment of the mixer is often required, (virtually instantaneous changes in the current product specifications). Three different meat + fat mixture series was studied directly in the IDE-CON mixer: bovine meat vs. fat (Figure 13), pork vs. fat (Figure 14) and (bovine + pork) vs. fat (Figure 15). The fat fraction to be added varied from 22% to 42 % in 4% steps in all three series.

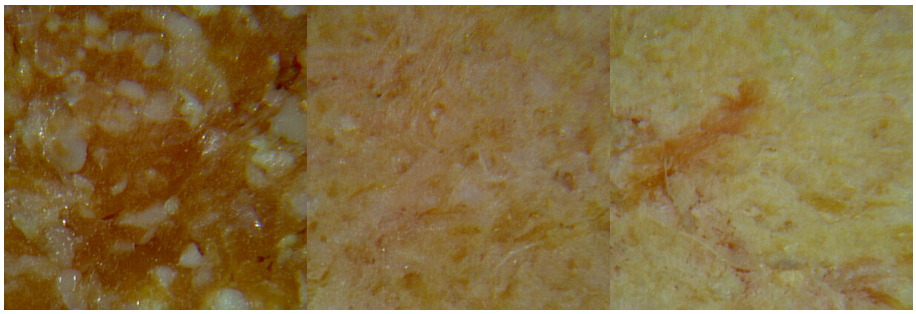


Figure 13. Bovine meat with incrementally added fat. From left to right: 21%, 33% and 41%. SILVACAM false colours.

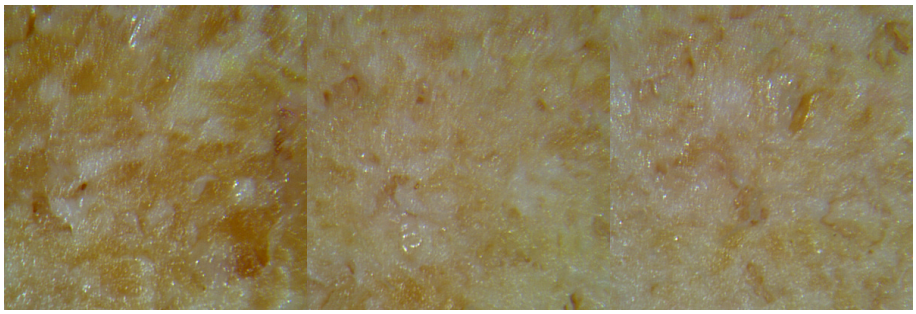


Figure 14. Pork with incrementally added fat fractions. From left to right: 22%,34% and 42% fat. SILVACAM false colours.

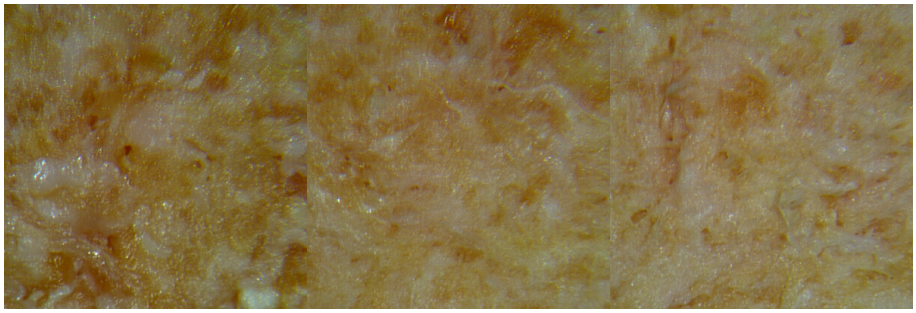


Figure 15. (Pork and bovine) with incrementally added fat. From left to right: 22%, 34%, and 42%. SILVACAM false colours.

Representative 150g samples were taken at the outlet of the IDE-CON mixer (full cross-sectional sampling) after identical mixing times (1 minute) for each new added fat-content increment of 4%. Samples were transported in glass petri-dishes, Figure 16, directly to our laboratory imaging setup presented below, Figure 17 with less than 20 minutes duration. There were no transportation segregation or similar in this type of mixture samples due to the extremely high viscosity of the meat-fat mixtures.

Figure 16. Petri-dishes with meat/fat mixtures. From top to bottom: Bovine, Bovine/Pork, Pork and pure fat. Fat content increases from left to right. Glass covers were removed just prior to imaging.

IMAGING SYSTEM

All studies reported here used a trusted, old-time friend of the Applied Chemometrics Research Group, the "SILVACAM" NIR/R/G digital camera (modified from an original JVC R/G/B television camera by the now defunct Finnish "Karelsilva" company (B. Braam). Figure 17 below presents the laboratory SILVACAM set-up.

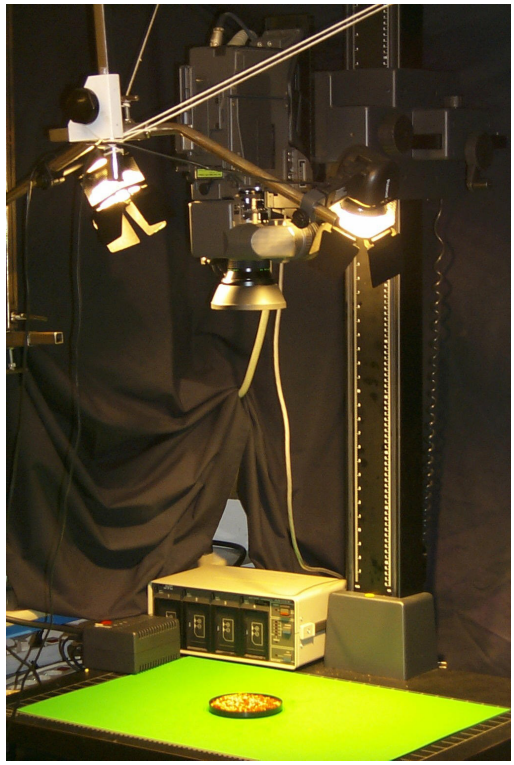


Figure 17. The SILVACAM imaging system at ACRG. Modified JVC-camera, with two quasi-parallel 150W illumination sources. Sampleholder (round) on sample table.

All image analysis systems are critically dependent upon a *proper* illumination system, which we have commented upon in several of our earlier powder and mixture studies[2-6]. Sometimes the AMT-derivation is directly dependent upon a unilateral low-angle illumination for example, while for other characterisations uniform multi-source illumination fits the bill. Each image analysis characterisation problem in fact always necessitates a thorough initial analysis of the proper illumination requirements for example. We shall here refrain from further commenting on this fundamental problem as all the examples used have been subjected to careful illumination optimisation efforts.

RESULTS

Results from applying the alternative AMT and MIR$^+$ approaches on the ten different data sets introduced above will be presented in parallel below. Focus is on *comparing* the optimal multivariate calibration models produced for each individual case; thus we performed *individual* multivariate calibration outlier screenings, and model-dimensional validations etc. for each model. For precisely this type of relative comparison purposes *full cross-validation* comes to its right with full force [7]. All models were calibrated against the pertinent mixing fraction as $\mathbf{y_1}$ variable (PLS1), while the relevant AMT-spectra or the alternative MIR$^+$-spectra served as the X-data block.

The number of objects (images) and replicates are equal in MIR$^+$ and AMT, but the pretreatment used wary in the examples. In some cases all replicates are shown (Figure 18 left), while in other cases replicate-spectra have been *averaged* (Figure 18 right), resulting in fewer objects in the model.

All AMT-spectra have been centered and scaled to uniform standard deviation (*auto-scaled*). In some MIR$^+$ cases, scaling the data can blow up noise and is thus avoided where possible. In other cases though, scaling the MIR$^+$ histograms was found a necessity.

Two-component Granular Mixtures

Black Pepper and White PVC

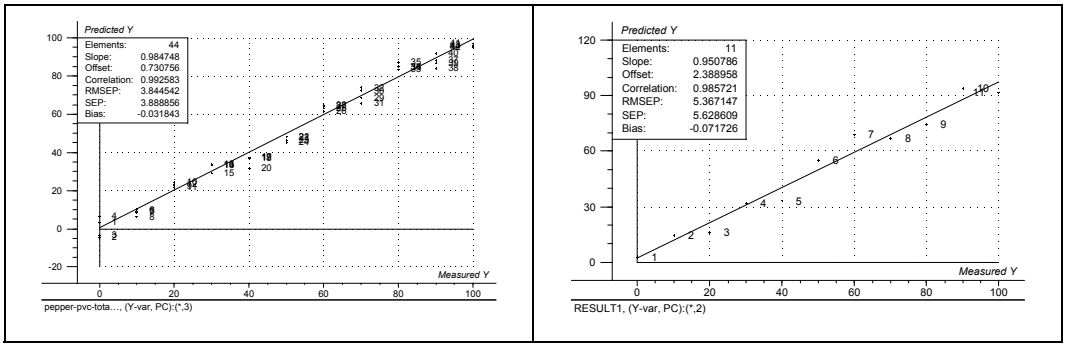


Figure 18. Black Pepper and PVC Pread-Meas plots. Left: MIR$^+$, Right: AMT

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	3	0.985	0.731	0.993	3.845

AMT	2	0.951	2.389	0.986	5.367

In the high-contrast ("easy") black pepper vs. white PVC-pellets case, MIR$^+$ performs *slightly* better with one additional component (fully significant according to the validation) is used.

This is one example where scaling is applied to the MIR$^+$ histogram. The predicted images are closely to a true black & white (one-bit) image, with information mainly in the beginning and end of the spectra. Scaling the data allows also the middle part of the histograms to influence on the model. Because of the very high contrast between the elements, the current example could possibly also be solved directly with thresholding.

Coriander & White Pepper

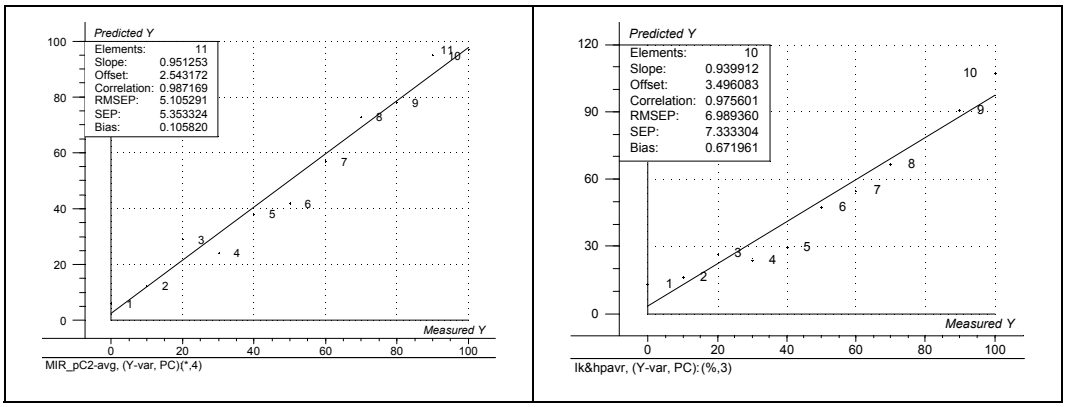


Figure 19. Coriander and White Pepper Pread-Meas plots. Left: MIR$^+$, Right: AMT

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	4	0.951	2.543	0.987	5.105
AMT	3	0.940	3.496	0.976	6.989

By using one more component, MIR$^+$ shows *marginally* better performance in the "very difficult" coriander vs. white pepper example.

In this example, scaling the MIR$^+$ data was not required. Because there is almost no contrast between the two classes, this example is not solvable using thresholding; a more subtle approach is required, hence MIR$^+$.

Grey and White PVC-pellets

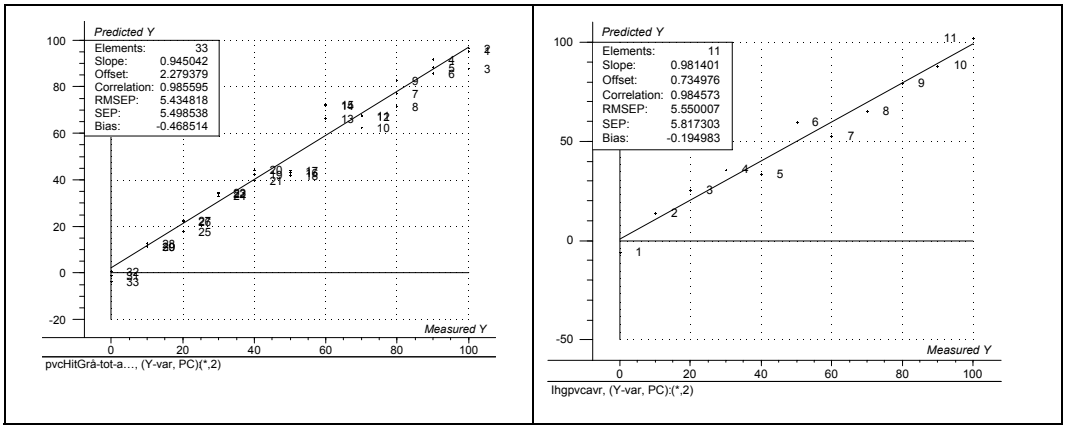


Figure 20. Grey and White Pread-Meas plots. Left: MIR^{+}, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR^{+}	2	0.945	2.279	0.986	5.435
AMT	2	0.981	0.735	0.985	5.550

In the case of the grey and white PVC-pellets the results are practically equal, although here AMT displays a clearly more comfortable slope (of a *fitted* "predicted vs. measured" regression).

This is another example of high spectral contrast, especially in the predicted images. Again, the data was scaled to extract information also from the middle parts of the histograms.

Grey PVC-pellets and green beans

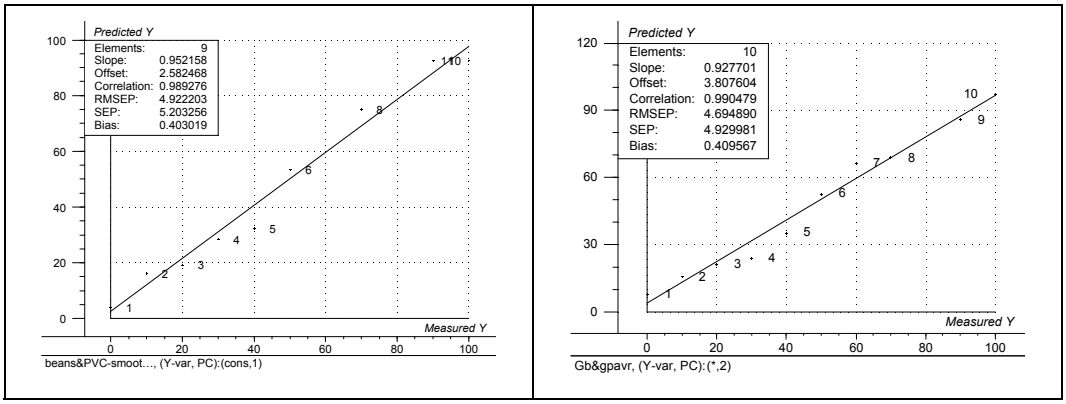


Figure 21. Grey PVC-pellets and green beans Pread-Meas plots. Left: MIR+, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR+	1	0.952	2.582	0.989	4.922
AMT	2	0.928	3.807	0.991	4.695

MIR+ has a slightly better performance in the grey PVC-pellets/green beans example, considering it uses less components.

In this example, the MIR+ histograms were not scaled. There is only small contrasts between the two elements, and the model uses the major "shape" of the histogram, and not so much the intermediate variables.

Three-Component Granular Mixture

PLS1 (y_1 : Peas mixing fraction)

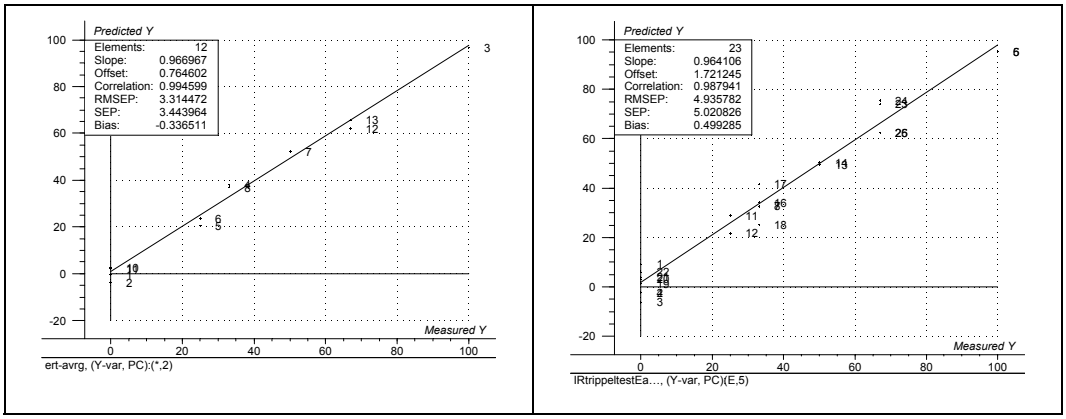


Figure 22. 3-component mixture modelled for Peas. Pread-Meas plots. Left: MIR$^+$, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	2	0.976	0.716	0.992	3.866
AMT	5	0.964	1.721	0.988	4.936

Concerning pea mixing fraction predictions, MIR$^+$ is clearly performing best, using two components vs. AMT which uses five.

In none of the three cases involving three-component mixtures, scaling were applied to the MIR$^+$ data. In the plots above (Figure 22), notice that the number of elements differ by a factor two. This is due to the use of a different averaging factor in the MIR$^+$ and AMT models. Final comparison is not hampered by this.

PLS1 (y_1 : Maize mixing fraction)

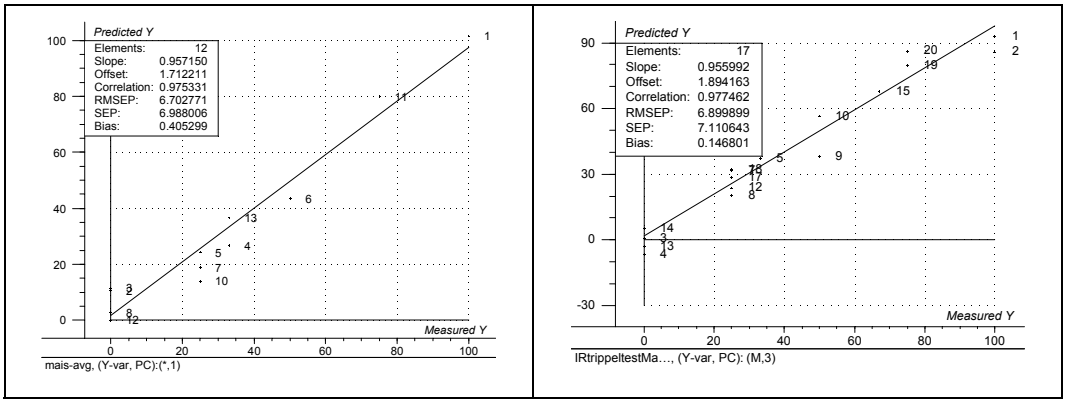


Figure 23. 3-component mixture modelled for Maize. Pread-Meas plots. Left: MIR$^+$, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	1	0.957	1.712	0.975	6.703
AMT	3	0.956	1.894	0.977	6.900

The maize prediction is difficult in both cases, and the results are almost identical. For maize both estimates of RMSEP are the largest of all three vegetables. The only difference is that MIR$^+$ uses one component, while AMT uses three.

PLS1 (y_1 : Carrot mixing fraction)

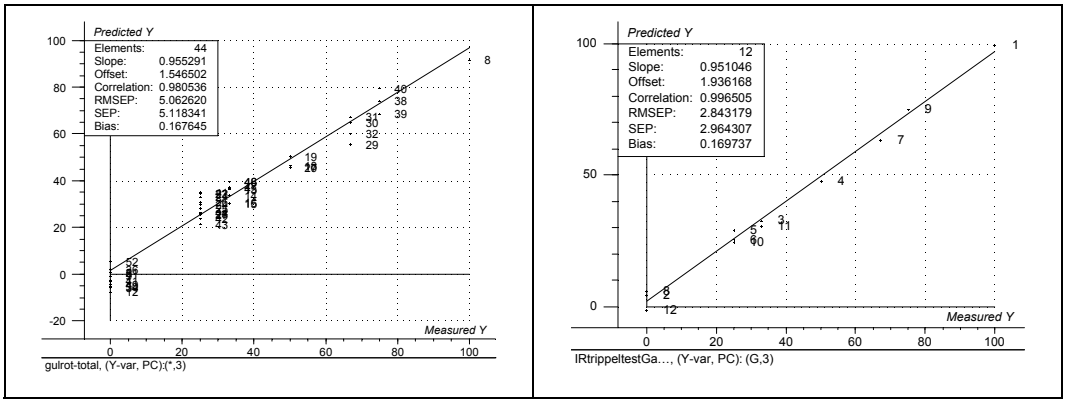


Figure 24. 3-component mixture modelled for Carrots. Pread-Meas plots. Left: MIR$^+$, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	3	0.955	1.546	0.981	5.063
AMT	3	0.951	1.936	0.997	2.843

In the carrot example AMT performs significantly better. The number of components are equal. Carrot cubes are clearly of a significantly different shape than either peas or maize.

Minced Meat

This example is organised in much the same way as the two above. MIR$^+$ training images, (Figure 5) were acquired for 100% pure meat of the relevant types and pure fat respectively, while in the experimental mixtures fat in the range of approx. 20-40% is studied, which is in the representative industrial production range. Meat was calibrated against a black Y-image, and fat calibrated against a white *ditto* (compare Figure 5). Starting at the reference minimum fat-content at 21%, the 4% fat increments (v/v) were added successively in a standardised manner and three replicate-samples were removed for each fat-level. For *each of these parallel physical replicates*, three images-replicates were also acquired by rotating the sample container 120 deg. in front of the camera. Thus there were a total of nine images representing each fat-level; there were overall six fat-levels in total, wiz. 21%, 25%, 29%, 33%, 37% and 41%.

Bovine

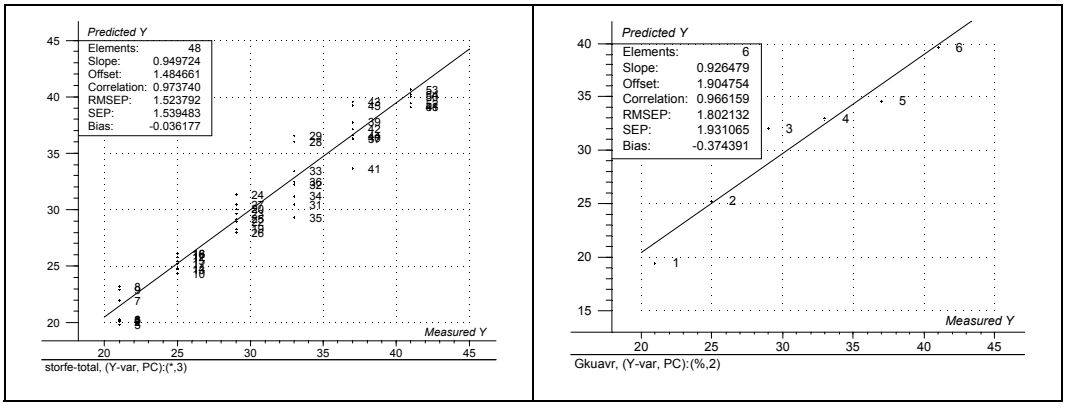


Figure 25. Minced Bovine meat and fat Pread-Meas plots. Left: MIR$^+$, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	3	0.950	1.485	0.974	1.524
AMT	2	0.926	1.905	0.966	1.802

MIR$^+$ predicts the fat content in Bovine meat slightly better, using one more PLS-component.

Pork

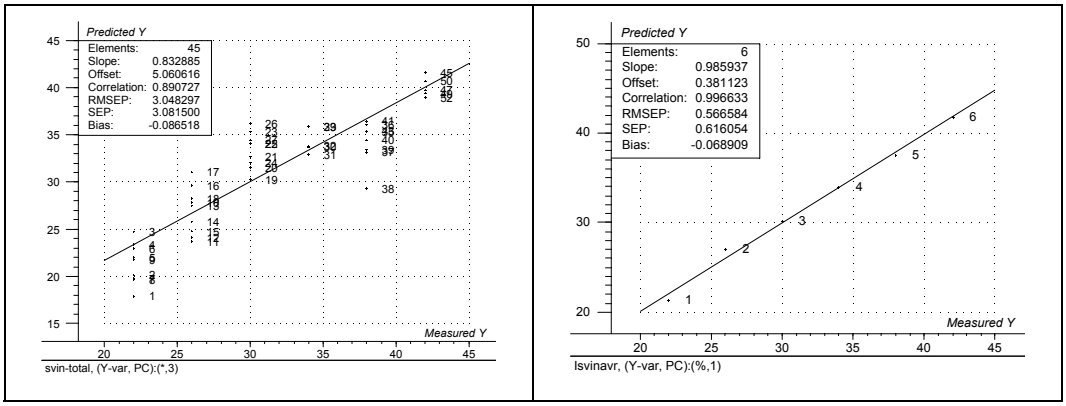


Figure 26. Minced Pork meat and fat Pread-Meas plots. Left: MIR$^+$, Right: AMT.

	# Comp	Slope	Offset	Correlation	RMSEP
MIR$^+$	3	0.833	5.061	0.891	3.048
AMT	2	0.986	0.381	0.997	0.566

In the pork example AMT performs significantly better, also boasting fewer PLS-components.

Bovine & Pork

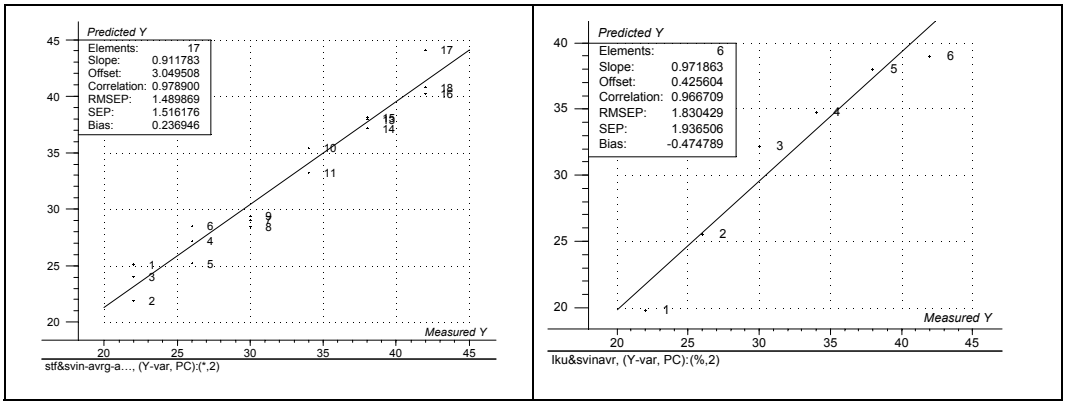


Figure 27. Minced Bovine & Pork meat and fat Pread-Meas plots. Left: MIR+, Right: AMT

	# Comp	Slope	Offset	Correlation	RMSEP
MIR+	2	0.912	3.050	0.979	1.490
AMT	2	0.972	0.426	0.967	1.830

In the combined meat (bovine and pork) vs. fat example, the results are practically equal, except w.r.t. the slope of the fitted regression index; this makes AMT a potential marginal winner here.

Discussion

Two-component mixtures - overview of results:

Essentially all models for both AMT and MIR+ perform satisfactorily in this first overview, but the MIR+ models do perform best, or marginally best, in three out of four detailed evaluations of the selected dry two-component mixture fraction prediction studies; one medium contrast ("somewhat difficult") case has AMT as best. It is encouraging that both the AMT - as well as the MIR+ models essentially both are up to the complicated image analysis job set up. The kinds of precisions obtained in these first attempts are satisfactory: For all models the total span of validation-estimated RMSEP ranges 3.845 - 7.866, while this range for the *best four models* corresponds to: 3.845 - 5.550. Based on an average mixing fraction of 50% these latter correspond to precisions of 7.7% - 11.1% respectively (rel. %). The better end of this interval comprise very respectable precisions for our first attempts image analysis approaches (irrespective of whether one chooses to improve on the MIR+ - or the AMT

approaches; a rational choice demands a much larger factual results data base than the one presently presented). It should be pointed out that the selected two-component mixing systems deliberately includes both the supposedly easiest - as well as the supposedly most difficult systems; also we were certainly surprised e.g. by the unexpected success of the coriander - white pepper case. The complete study of representative dry two-component mixing systems is far from finished at present, and shall be reported on in its totality at a later occasion. The present results can only be characterised as very *encouraging* as feasibility studies go.

Three-component vegetable mixtures:

Winners are about equal, wiz. one AMT- and one MIR$^+$ model, and one draw (maize). Validation estimates of RMSEP are - peas: 3.866 (MIR$^+$); maize: 6.703 (MIR$^+$/AMT); carrots: 2.843 (AMT), which translates to the following rel. % precision (+/- 1 RMSEP) - peas: 7.7%; maize: 13.4% and carrots: 5.7% respectively (all calculated w.r.t. an average mixing fraction of 50% (abs.). Two of these three models, characterised by very realistic sample preparation variances, actually reach *below* industry's precision demand of 8% already from these *first pilot studies* (sic). Clearly the troublesome maize prediction can be better handled by a simple *constant sum* difference calculation! We term these pilot results as *absolutely satisfactory*.

Minced meat mixtures:

For the three best minced meat models, the validation RMSEP estimates translates to 4.9% (MIR$^+$), 1.8% (AMT) and 5.9% (AMT) respectively (all expressed as relative %), compared to a product fat specification range of 21-41% (rel. % calculated w.r.t. an average of 31%). For a first pilot study of this relatively complex on-line mixing system, precisions of 2-6 rel. % can only be characterised as *excellent.* Not only are the outlet sampling procedures not fully optimised yet, neither are the imaging illumination conditions etc. At this time it is only possible to say that there is certainly a significant *potential* improvement to be gained here.

AMT or MIR$^+$:

If judgement would have to passed on the basis of the present results alone, the new, extended MIR$^+$ approach merits very close attention. The degree of accuracy and precision obtained for the present *three very different sets of mixing systems* is impressing indeed. And AMT is a very close runners-up, which should also be related to its recent history of well-documented successes, reported in several complementary

powder/granular matter studies [2-5], in which AMT has shown similar degrees of promising results. Thus we are not at all in a(ny) position to even try to chose between these two powerful chemometric image analysis processing alternatives - on the contrary: they will of course be extensively further explored in parallel.

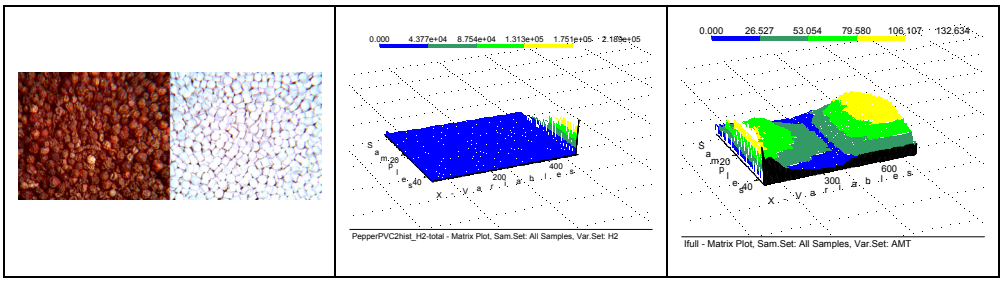
The above discussion has presented more than enough evidence to leave this pilot study in a very optimistic state. Much interesting real-world, industrial implementation work awaits, as well as an enticing fundamental science, broad-scoped functional powder mixing laboratory study - all with exciting potentials for industrial applications and economic benefits.

Interpretations of alternative model structures: MIR$^+$ vs. AMT

We end this study by showing what kind of detailed interpretations is made possible by systematic evaluation of the prediction models established. By focusing on the final *validated* prediction results, we now examine the models from the perspective of their comparative *intrinsic data structures*, as represented by their loading-weight relationships.

At the same time we shall also honour the inter-comparison objective between AMT and MIR$^+$, although the conclusions above would appear to rule out any unambiguous winner. Thus this comparison shall mostly focus on the basic differences experienced in their respective w-relationships. These two fundamentally different approaches to quantification of mixture – and texture features will necessarily lead to very different loading-weight spectra (they are *de facto* modelling quite different features) – while in the final assessment they may well lead to almost identical prediction strengths, as indeed demonstrated above.

Figure 28 shows the necessary interpretation background in the form of a small recap of the raw images and their corresponding MIR$^+$ and AMT raw spectra respectively. This juxtaposition will make it easy to appreciate the following interpretations.



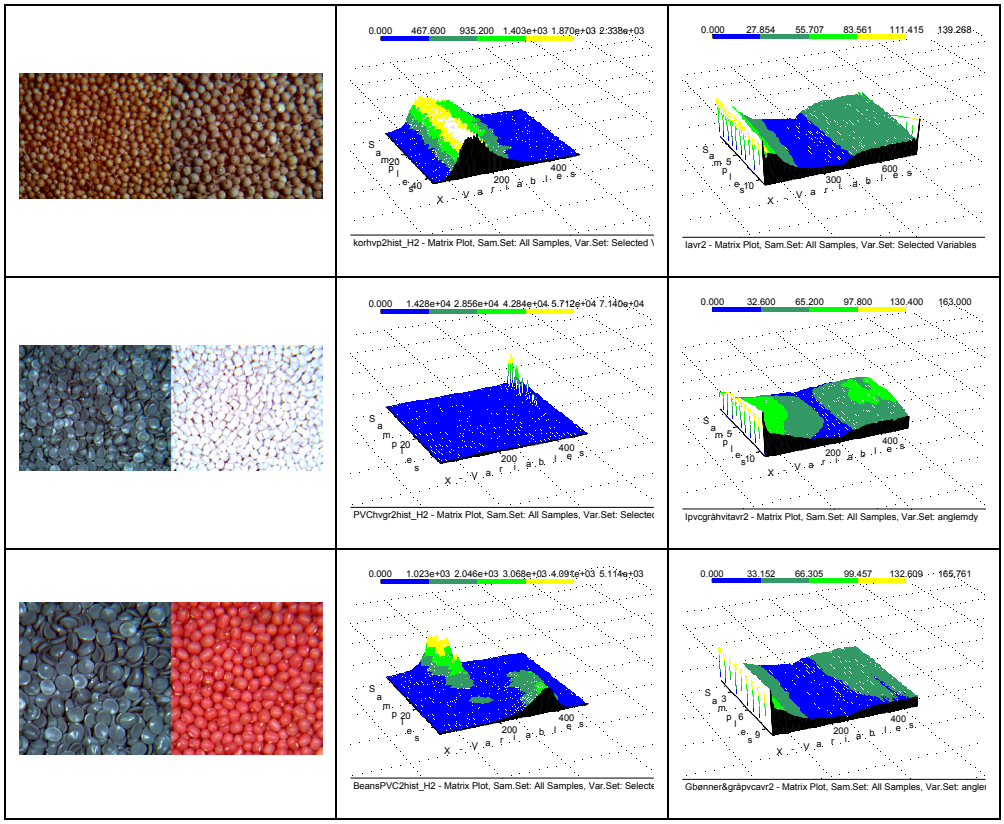
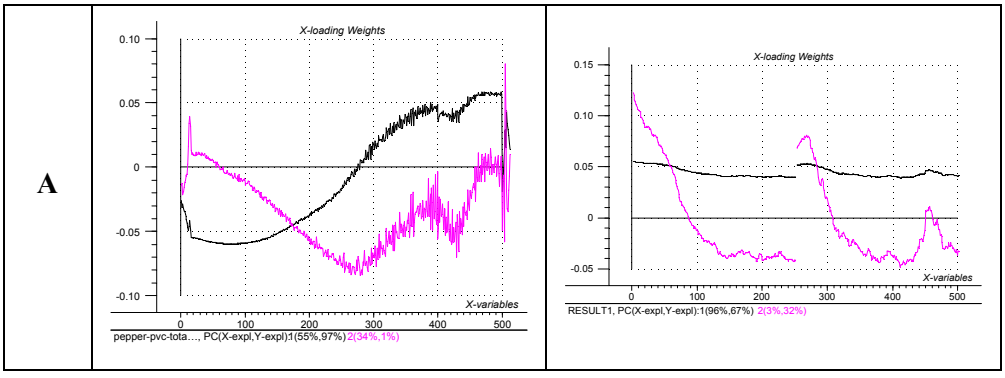


Figure 28. Data presentation of raw data (left), MIR^{+} histograms (middle) and AMT-spectra (right).

Figure 28 shows a two-fold division in high-contrast (row A and C) and intermediate-contrast mixing systems (row B and D). Individual MIR^{+}-modelling lead to the use of *auto-scaled* models for the former, while the latter were best serviced without. One observes the very marked different raw MIR^{+}-spectra for these opposing systems.

These observations makes for easy detection of a similar pattern in both the MIR^{+} and the AMT w-spectra below in Figure 29.



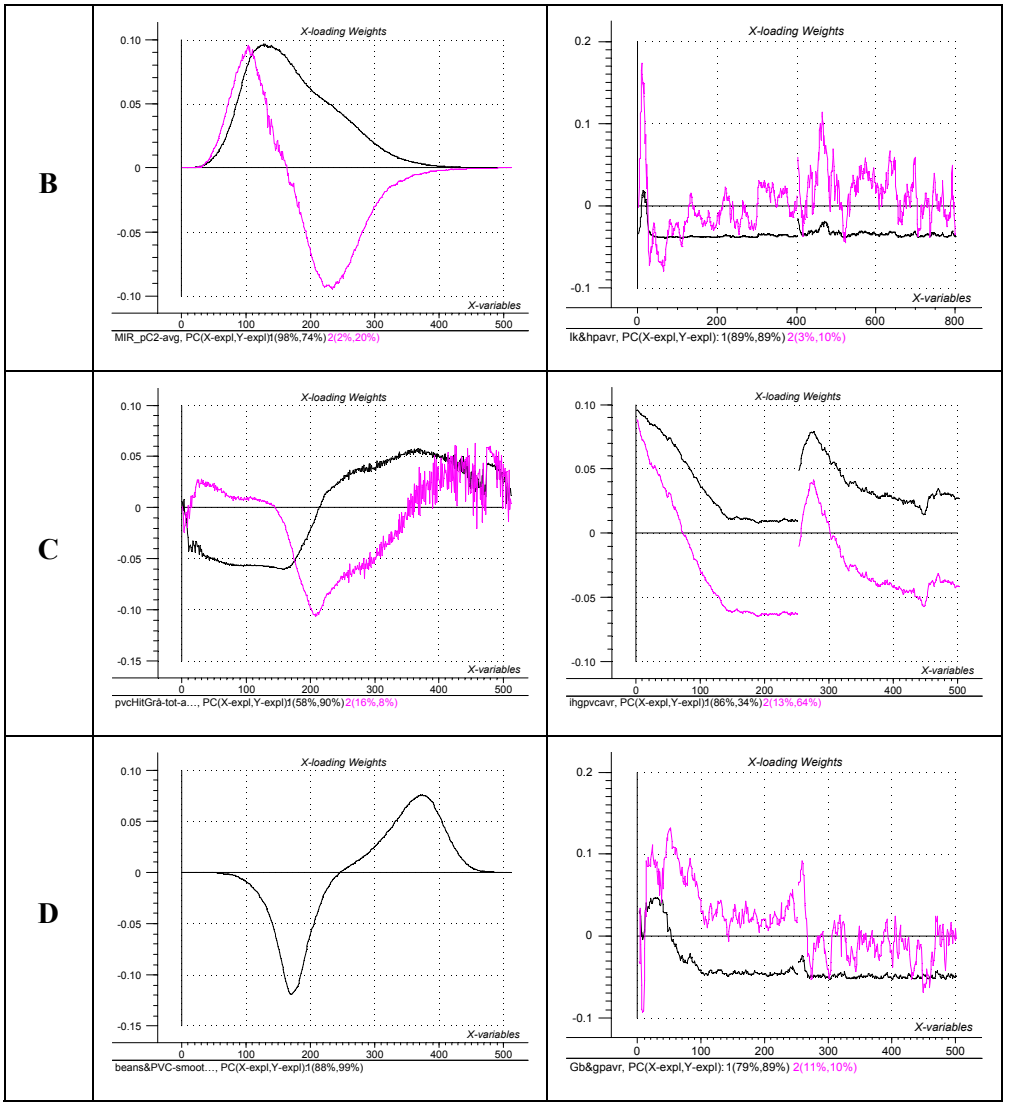


Figure 29. 2-component mixture loading weight plots. Left: Mir$^+$, right: AMT. From top to bottom: Pepper/PVC, Coriander/White Pepper, white/grey PVC and Beans/PVC.

MIR$^+$:

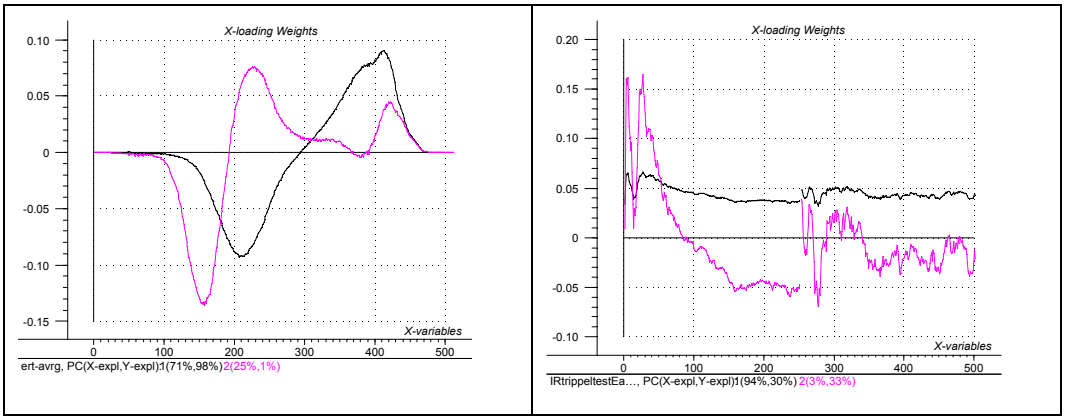
MIR$^+$-modelling of the *intermediate-contrast systems* are distinctive (row B and D). The first PLS-component, to a large extent, takes good care of the Y-modelling but with a very significant *addition* from PLS-component 2 (row B). For both systems w_1 mimics the raw MIR$^+$-spectrum to a very high degree, while the second order *addendum* from PLS-component 2 attest to a slight shift in the X-variable direction for the coriander-white pepper system (row B); for row D system there is an even simpler relationship with mixing fractions (Y) leading to only one significant PLS-component.

For the more-simple-to-model *high-contrast systems* (row A and C) one observes that two PLS-components are also needed, but with only marginally improved Y-variances for the second PLS-component (only first significant according to the validations). It will be appreciated that the X-variance is utilised in a very effective fashion caused by the auto-scaling.

AMT:

For AMT one observes a distinctly opposite pattern. For both high-contrast systems (row A and C) there is now a very marked need for both PLS-components in order to do the prediction modelling effectively (32% and 64% Y-variance accounted for respectively by the second component). Both components are now *highly significant* according to the validations. For both systems, their respective w_1- and w_2-spectra shows essentially the same pattern (rare!), while for the opposing intermediate-contrast systems (row B and D) there is only a small (10%) Y-variance addendum from w_2.

With due reference to the rather disparate four systems some underlying systematics may perhaps be found. MIR$^+$ manages to combine most of the essential X-variance in just one PLS-component (three out of four systems), while AMT would appear to favour two-component systems, especially for the high-contrast cases. Our initial classification into H, I and L-contrast systems may very well be further refined a.o. also based upon this kind of systematic modelling of all systems in the background study (nine systems covering the H, I, L-domain more fully).



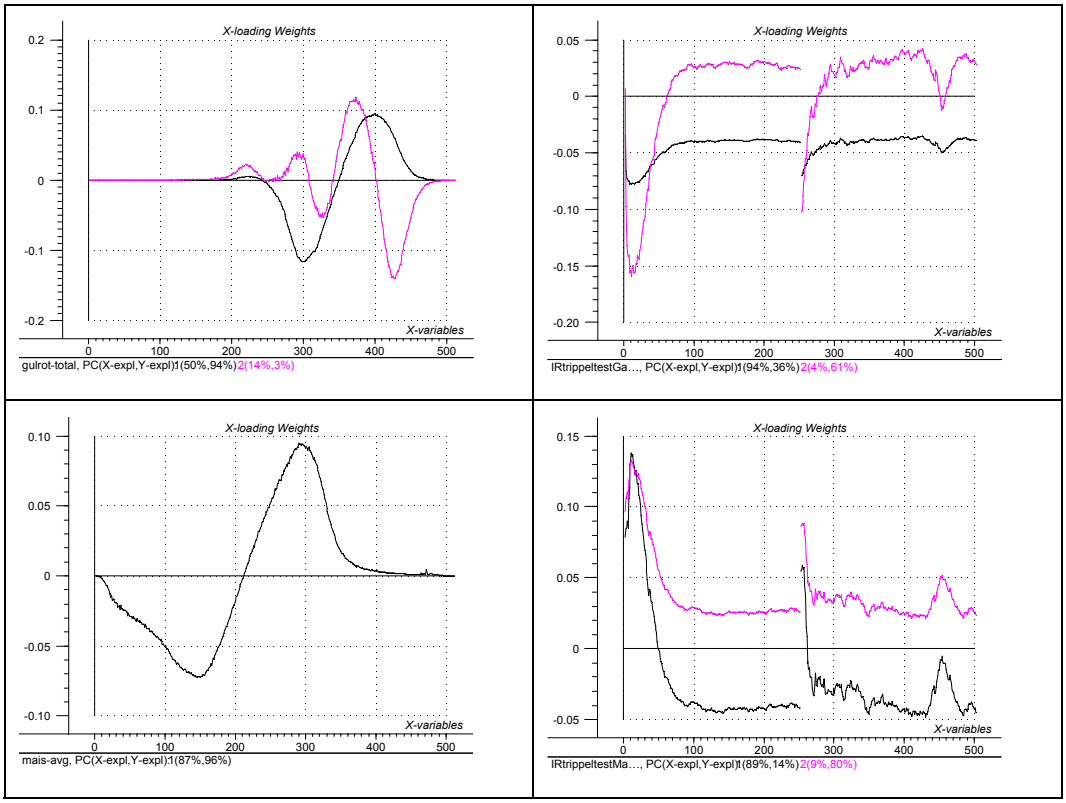


Figure 30. 3-component loading weights plots. Left: MIR$^+$, right: AMT. From top to bottom: Peas, Carrot and Maize.

For the three-component vegetable system we may use the above interpretation systematics in order to simplify what would *appear* to be a more complex issue.

Inspection of Figure 30 again reveals an extremely simple relationship for MIR$^+$ however. For both peas, carrot as well as maize prediction models, the first PLS-component accounts for 98%, 96% and 94% Y-variance respectively, with barely significant, very minor additions for the second components, a very clear one-component trend.

In stark contrast to this, the AMT-relationships show marked multi-component features, some using even more than two validated components, thus further contributing to the overall MIR$^+$ vs. AMT relative pattern. MIR$^+$ is able to model even these, clearly more complex systems, still basically using only one PLS-component – no doubt primarily due to its underlying dichotomous 0/100 model-definition. AMT on the other hand, while able to reach essentially identical prediction validation results, does this in a distinctly more elaborate fashion in which several essential *contrast phenomena* are found distributed over more PLS-components.

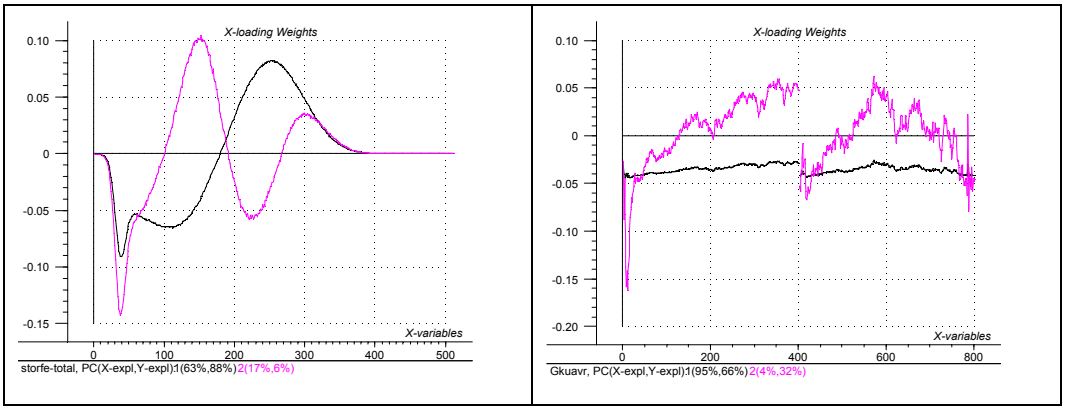
The remaining meat-fat systems (Figure 31 below) are no doubt one order of magnitude more complex still. But here again for MIR$^+$ we find the exact same dispositions as for the vegetable – and the dry powder systems both: extreme reliance on the first PLS-component accounting for 88%, 98% and 93% Y-variance respectively, while AMT here shows especially complex multi-component patterns, also between the different meat-fat mixing series internally. In fact there would appear to be very interesting detailed interpretation possibilities for this latter system as regards these internal AMT differences, which we shall never-the-less leave for an other occasion since the overall *comparative* MIR$^+$ vs. AMT pattern remains the same for this system as for the two above:

Considering the gamut of all the three pilot studies, covering a broad swath of relevant real-world, industrial mixing two-component and three-component end-member systems, the overall conclusion would now appear to have become clear:

MIR$^+$ can do with few – what AMT must do with more.

Following Occam's razor, we then must point to MIR$^+$ as a very powerful new complement to AMT in the family of multivariate image regression problem-dependent pre-processing facilities, which we intend to develop much further with great interest.

AMT on the other hand, confirms its status of being able to model even very complex systems with a detailed internal model structure, well suited for in-depth interpretations.



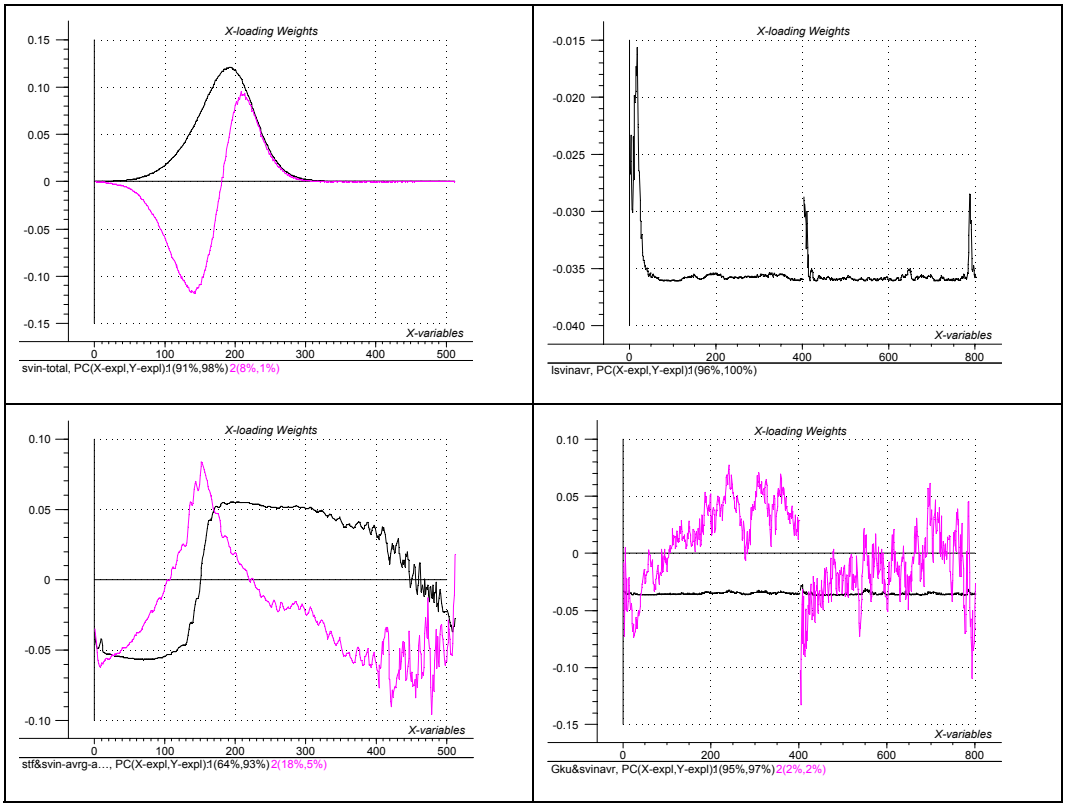


Figure 31. Minced Meat loading weight plots. Left: MIR+, right: AMT. From top to bottom: Bovine, Pork and Bovine+Pork.

ACKNOWLEDGEMENTS

Ide-Con was founded 1989 and has been granted three patents on new mixing concepts. The mixer has been tested out at the Powder of Science and Tecnology (POSTEC), Porsgrunn Norway as well as in several key industries. We are very greatful to Geir Nordahl and Thor Ragnar Hem for permission to publish results here based on our first collaboration using IC-B13.

REFERENCES

[1] T. Shinbrot and F. J. Muzzio, *Nonequilibrium Patterns in Granular Mixing and Segregation.* Physics today, **53/3**. (2000) pp 25-30

[2] J. Huang and K.H. Esbensen, *Applications of AMT (Angle Measure Technique) in Image Analysis Part I. A New Methodology for in-situ Powder Characterization.* Chemometrics and Intelligent Laboratory Systems. **54/1**. (2000) pp 1-19

[3] Jun Huang and Kim H. Esbensen . *Applications of AMT (Angle Measure Technique) in Image Analysis PART II: Prediction of Powder Functional Properties and Mixing Components using Multivariate AMT Regression (MAR).*, accepted by Chemo. Lab. 2000.

[4] J. Huang, B. Møller, K.H. Esbensen, L. Munck, *Characterization of Barley Germination Process from the Images by Combined AMT(Angle Measure Technique) Chemometric Analysis*. Submitted for publication in Analytica Chimica Acta, 2000

[5] K.H. Esbensen and J. Huang, *High sensitivity particulate impurity detection and quality control by unfolded image AMT.* Submitted for publication in Journal of Chemometrics, 2000

[6] J. Huang and K.H. Esbensen, *Methodology Development and Applications of Multi-way methods in Image Analysis.* Submitted for publication in Journal of Chemometrics, 2000

[7] K. Esbensen et.al. *Multivariate Data Analysis –in practice, 4th edition.* (2000) CAMO ASA, Oslo, Norway. ISBN 82-993330-2-4.

[8] P. Geladi and K. Esbensen. *Regression on multivariate images: Principal Component Regression for modelling, prediction and visual diagnostic tools.* J. of Chemometrics **5** (1991) 97-111.

[9] K. Esbensen, P. Geladi and Hans Grahn. *Strategies for Multivariate Image Regression.* Chemo. Lab. **14** (1992) 357-374.

[10] P. Geladi and H. Grahn. *Multivariate Image Analysis.* (1996). John Wiley and Sons, Chichester, UK pp. 316

[11] F. Lindgren, P. Geladi and S. Wold. *The Kernel Algorithm for PLS.* J. of Chemometrics, **7** (1993) 45-59.

[12] S. Wold, H. Martens and H. Wold. *The multivariate Calibration problem in chemistry solved by the PLS method*. (1983) Proc. Conf. Matrix pencils, (A. Ruhe, B. Kågstrøm, eds), March 1982. *Lecture Notes in Mathematics,* Springer Verlag, Heidelberg, 286-293

[13] H. Martens and S.Å. Jensen. *Partial Least Squares regression: A new two-stage NIR calibration method.* (1983) Proc. 7th World Cereal and Bread Congress. Prague June 1982. (Holas and Kratochvil, eds.) Elsevier Publ., Amsterdam, 607-647.

[14] H. Martens and T. Næs. *Multivariate Calibration.* (1989). John Wiley & Sons Ltd. Chichester, UK. ISBN 0 471 90979 3 pp. 419

[15] T.T. Lied, P. Geladi and K. Esbensen. *Multivariate Image Regression (MIR): implementation of image PLSR – first forays*. J. Chemometrics, **14** (2000) pp. 585-598

[16] T.T. Lied and K. Esbensen. Principles of MIR, Multivariate Image Regression I: regression typology and representative application studies. (Submitted for publication).

[17] T.T. Lied and K. Esbensen. *Principles of MIR, Multivariate Image Regression II: Cross Validation – what you see is what you get.* (Submitted for publication, 2000).

[18] R. C. Gonzalez and R. E. Woods. *Digital Image Processing.* (1993) Addison-Wesley Publishing Company, Inc. Reading, USA. ISBN 0-201-60078-1 pp. 716.